



Proyecto Fin de Carrera
Ingeniería Informática

**FOTOGRAFÍA COMPUTACIONAL:
REPRODUCCIÓN DE TONO INVERSA PARA
IMÁGENES SUBEXPUESTAS Y
SOBREEXPUESTAS**

Miguel Martín González

Director: Dr. Diego Gutiérrez Pérez

Departamento de Informática e Ingeniería de Sistemas
Centro Politécnico Superior
Universidad de Zaragoza

Junio 2008

Índice general

Índice de figuras	4
Índice de tablas	7
1. Objeto y alcance	9
1.1. Introducción	9
1.2. Objetivos	11
1.3. Alcance	12
1.4. Estructura de la memoria	14
1.4.1. Capítulo 2	14
1.4.2. Capítulo 3	15
1.4.3. Capítulo 4	15
1.4.4. Capítulo 5	15
1.4.5. Capítulo 6	15
1.4.6. Apéndice A	15
1.4.7. Apéndice B	15
1.4.8. Apéndice C	15
1.4.9. Apéndice D	15
1.4.10. Apéndice E	16
2. El problema	17
2.1. Definiendo subexposición y sobreexposición	17
2.1.1. Introducción	17
2.1.2. Clasificación de imágenes: tests psicofísicos	20
2.2. Diseñando un nuevo rTMO	21
2.2.1. Comparación de rTMO's existentes	21
2.2.2. Adecuación de cada ITMO en función de la exposición	21
2.2.3. Diseño del algoritmo	22
3. Clasificando las imágenes: test psicofísico 1	23
3.1. Diseño del test	23
3.1.1. Motivación	23
3.1.2. Estímulos	24
3.1.3. Metodología	25
3.2. Resultados y conclusiones	27

3.3. Conclusiones	30
4. Comparando rTMO's: test psicofísico 2	34
4.1. Estado del arte: principales rTMO's	35
4.1.1. <i>Linear expansion</i>	35
4.1.2. <i>Inverse Tone Mapping algorithm</i> [BLDC06], por <i>University of Bristol</i>	35
4.1.3. <i>The reproduction of Specular Highlights on High Dynamic Range Displays</i> [MDS06], por <i>Ecole Polytechnique Fédérale de Lausanne, Switzerland</i> y <i>Sharp Laboratories of America, Camas WA, USA</i> . .	37
4.1.4. <i>Ldr2hdr: On-the-fly Reverse Tone Mapping</i> [RTS ⁺ 07], por la <i>University of British Columbia</i> y <i>Dolby Canada</i>	39
4.2. Diseño del test	40
4.2.1. Estímulos	40
4.2.2. Metodología	41
4.3. Resultados	44
4.3.1. Mejor rTMO global	44
4.3.2. Mejor rTMO para subexposición	44
4.3.3. Mejor rTMO para sobreexposición	45
4.3.4. Estadísticas de la imagen vs rTMO	46
4.4. Conclusiones	47
5. Diseñando un nuevo rTMO	50
5.1. Análisis de la fotografía LDR a transformar	50
5.2. Aplicación del rTMO óptimo	51
5.3. Normalización de resultados	52
5.4. Conclusiones	52
6. Problemas, conclusiones y trabajos futuros	53
6.1. Conclusiones y trabajo futuro	53
6.2. Problemas encontrados	55
A. Conceptos fotográficos	56
A.1. Introducción	56
A.2. Exposición	57
A.3. Histograma	61
A.4. Soluciones clásicas al problema de la limitación del rango dinámico.	62
A.5. El rango dinámico en la fotografía clásica	62
A.6. El rango dinámico en la fotografía en color	63
B. HDR	65
B.1. Introducción	65
B.2. Captura	66
B.3. Almacenamiento	67
B.3.1. PFM (<i>Portable Float Map</i>)	68
B.4. Edición	69

B.5. Tone Mapping (mapeado de tono)	70
B.6. Visualización	71
B.7. Aplicaciones	74
B.7.1. Películas	75
B.7.2. Juegos	75
B.7.3. Medicina	75
B.7.4. Datos geológicos	76
B.7.5. Imágenes por satélite	77
B.7.6. Nuevos algoritmos	77
C. Psicofísica	78
C.1. Psicofísica: conceptos	78
C.2. Prácticas empleadas	79
C.2.1. Periodo de <i>entrenamiento</i>	79
C.2.2. Aleatorización de muestras	79
C.2.3. Escala de medición conocida	80
C.2.4. Público heterogéneo	80
C.2.5. Elementos de control	81
C.2.6. Libertad de modificación	81
D. Estímulos	82
E. Artículo de investigación presentado al CEIG2008	88
Bibliografía	89

Índice de figuras

1.1. Obtención de imagen HDR a partir de horquillado de exposición. (Stanford Memorial Church, fotografía de Paul Debevec [Deb])	10
1.2. Demostración del efecto de la exposición en la fotografía nocturna	11
1.3. Idea del algoritmo-objetivo final.	12
2.1. Imágen sobreexpuesta e imagen correcta	18
2.2. Misma fotografía: zonas correctamente expuestas y zonas sobreexpuestas	18
2.3. Subexposición y sobreexposición en la misma fotografía	19
2.4. Fotografías que comparten histogramas, pero no corrección en la exposición	20
3.1. Estímulos: un representante de cada escena.	24
3.2. La secuencia completa de bracketing para la escena <i>sunset</i> , junto a sus respectivos histogramas de luminancias.	25
3.3. Interfaz usuario test 1	28
3.4. Resultados numéricos del test psicofísico	29
3.5. Resultados del test psicofísico: taxonomía de los estímulos utilizados	32
3.6. Correlación entre los resultados del test psicofísico para subexposición y la media de luminancia de los pixels de las imágenes	32
3.7. Correlación entre los resultados del test psicofísico para sobreexposición y el porcentaje de pixels sobreexpuestos	33
4.1. Proceso de expansión lineal.	35
4.2. <i>Inverse Tone Mapping</i> : creación del <i>expand map</i>	36
4.3. El rTMO de <i>Ecole Polytechnique Fédérale de Lausanne</i>	37
4.4. Variación de ρ en el rTMO de <i>Lausanne</i>	38
4.5. <i>Ldr2hdr: On-the-fly Reverse Tone Mapping</i> : proceso	39
4.6. Algoritmo de <i>edge stopping</i>	40
4.7. Estímulos: un representante de cada escena.	41
4.8. La secuencia completa de bracketing para la escena <i>sunset</i> , junto a sus respectivos histogramas de luminancias.	41
4.9. Comparando rTMO's: <i>LDR</i> vs <i>Linear</i> vs <i>Warwick</i> vs <i>Dolby</i>	43
4.10. Resultados del test psicofísico: Comparativa global entre rTMO's.	45
4.11. Resultados del test psicofísico: Comparativa entre rTMO's para subexposición.	45

4.12. Resultados del test psicofísico: Comparativa entre rTMO's para sobreexposición.	46
4.13. Resultados del test psicofísico: Comparativa entre rTMO's en imágenes con media lumínica baja.	47
4.14. Estímulos percibidos como sobreexpuestos	48
4.15. Resultados del test psicofísico: Comparativa entre rTMO's en imágenes con alto porcentaje de pixels saturados.	49
5.1. Pasos del nuevo rTMO: análisis, aplicación de rTMO en función al análisis y normalización de salida.	52
A.1. Limitación del rango lumínico admisible por el medio.	57
A.2. Bracketting: variación de la exposición.	58
A.3. Rango visible de valores para fotografía tradicional y fotografía HDR.	58
A.4. Efecto de profundidad de campo	59
A.5. Apertura y profundidad de campo	59
A.6. Modo fotográfico de prioridad a la velocidad	60
A.7. Exposición automática (usando Sony CyberShot DSC-P100)	60
A.8. Fotografía a contraluz (tomada con Panasonic Lumix FZ-8)	61
A.9. Tipificación de la imagen usando el histograma	62
A.10. La agonía en el Jardín, de El Greco	63
A.11. Capas del negativo a color	64
A.12. Usando filtros de color: obra de Prokudin Gorskii en 1915	64
B.1. Proceso fotográfico: captura, almacenamiento, edición y salida	65
B.2. Sensores en chips fotográficos	66
B.3. Muestra de algunos píxeles antes y después de la normalización presentada porDebevec en SIGGRAPH97	67
B.4. Salida de XVI32 para archivo PFM	69
B.5. Combinando varias imágenes LDR para obtener una HDR, usando Adobe Photoshop.	70
B.6. Taxonomía espacio-temporal de los operadores de tono	71
B.7. Sistema de visualización BrightSide DR37-P	72
B.8. Corte de un monitor de nueva generación.	73
B.9. Comparativa entre distintos rangos lumínicos.	73
B.10. Funcionamiento de los monitores HDR	74
B.11. Monitor HDR vs monitor convencional	74
B.12. Monitor HDR vs monitor convencional (2)	74
B.13. Monitor HDR vs monitor convencional (3)	75
B.14. Usos médicos de HDR (1): rayos X	76
B.15. Usos médicos de HDR (2): laparoscopia	76
B.16. Usos geológicos de HDR: oceanografía	76
D.1. Secuencia 01: <i>building</i>	83
D.2. Secuencia 02: <i>car</i>	83
D.3. Secuencia 03: <i>indoor</i>	84

D.4. Secuencia 04: <i>lake</i>	84
D.5. Secuencia 05: <i>pencils</i>	85
D.6. Secuencia 06: <i>computers</i>	85
D.7. Secuencia 07: <i>waxes</i>	86
D.8. Secuencia 08: <i>sunset</i>	86
D.9. Secuencia 09: <i>graffitti</i>	87
D.10. Secuencia 10: <i>strawberries</i>	87

Índice de tablas

3.1. Luminancia media de los píxeles para cada una de las fotografías de cada secuencia.	25
3.2. Mediana de luminancia de los píxeles para cada una de las fotografías de cada secuencia.	26
3.3. Porcentaje de pixels con valor lumínico 254 y 255.	26
3.4. Porcentaje de pixels con valor lumínico nulo.	27
C.1. Escala BorgCR10	80
D.1. Formato de datos estadísticos de los estímulos	82
D.2. Datos estadísticos de los estímulos secuencia <i>building</i>	83
D.3. Datos estadísticos de los estímulos secuencia <i>car</i>	83
D.4. Datos estadísticos de los estímulos secuencia <i>indoor</i>	84
D.5. Datos estadísticos de los estímulos secuencia <i>lake</i>	84
D.6. Datos estadísticos de los estímulos secuencia <i>pencils</i>	85
D.7. Datos estadísticos de los estímulos secuencia <i>computers</i>	85
D.8. Datos estadísticos de los estímulos secuencia <i>waxes</i>	86
D.9. Datos estadísticos de los estímulos secuencia <i>sunset</i>	86
D.10. Datos estadísticos de los estímulos secuencia <i>graffitti</i>	87
D.11. Datos estadísticos de los estímulos secuencia <i>strawberries</i>	87

Agradecimientos

Me siento afortunado de haber trabajado con personas estupendas. En primer lugar, me gustaría expresar mi mayor agradecimiento a Diego Gutiérrez, por haberme dado la oportunidad de llevar a cabo este proyecto, así como por sus comentarios, sugerencias, apoyo y ánimos. Me complace, además, haber tenido la oportunidad de observar *in situ* el BrightSide DR37-P. Desde aquí mi más profundo agradecimiento a Karol Myszkowski y Matthias Ihrke (Max Planck Institute for Informatics) por sus atenciones y consejos durante mi estancia en Saarbrucken (Alemania).

Gracias a Choss y al resto de fotógrafos que han contribuido en este Proyecto. Su experiencia y consejos han sido determinantes.

Asimismo ha sido indispensable la colaboración de Francesco Banterle y Olga Sorkine, que han dedicado tiempo y esfuerzos a aplicar sus algoritmos a nuestras imágenes. Gracias también a Roland Fleming por sus consejos a la hora de diseñar los tests psicofísicos, así como por las horas de su tiempo invertidas a supervisar los experimentos. Del mismo modo, me gustaría agradecer a los participantes de los tests psicofísicos, por su paciente colaboración, y porque sin ellos nada de esto habría sido posible.

Gracias de nuevo a Olga, Roland y Diego, co-autores del paper *Understanding overexposure for reverse tone mapping*, (presentado para el CEIG2008 y pendiente de aprobación a fecha de cierre de este proyecto) por las sugerencias y esfuerzos invertidos.

Y sobre todo, gracias a mi familia, mi novia y mis amigos, por la cantidad de horas que han *sufrido* escuchando comentarios de un proyecto que no les concierne, y por sus ánimos y apoyo en los momentos difíciles. Sin ellos no habría llegado hasta aquí.

Capítulo 1

Objeto y alcance

1.1. Introducción

Hoy en día, las imágenes de alto rango dinámico (High Dynamic Range Imaging, o HDRI) reciben una cuidadosa atención desde el punto de vista cinematográfico e industrial. El margen dinámico es el rango de iluminación (EV) en el que la cámara es capaz de ver detalle. Un margen dinámico elevado permite que, pese a que la instantánea tenga zonas con luces fuertes y sombras oscuras, se mantenga el detalle en ambas iluminaciones. Cuando se supera el margen dinámico, las zonas sobreexpuestas¹ aparecen totalmente blancas y las sombras totalmente negras, de manera que resulta imposible recuperar la textura original.

En los gráficos por ordenador y en la fotografía, HDRI es un conjunto de técnicas que permiten un mayor rango dinámico de exposición que las técnicas usuales de fotografía digital. La intención de la tecnología HDRI es representar de forma precisa el amplio rango de niveles de intensidad que se encuentran en las escenas reales, desde la luz directa del sol a las sombras más oscuras. HDRI permite el almacenamiento de un amplio abanico de valores fotométricos físicamente precisos para cada píxel, asemejándose a los que es capaz de percibir el sistema visual humano [RWPD05].

HDRI se desarrolló originalmente para ser utilizado de forma exclusiva con gráficos generados por ordenador. Después, estos métodos se aplicaron para producir imágenes de alto rango dinámico a partir de una serie de fotografías tomadas con diferentes exposiciones (técnica conocida como horquillado de exposición o *exposure bracketing*) [DM97].

Debido a la creciente popularidad de las cámaras digitales, y a la facilidad de manejo del software existente, el término HDR se usa de forma popular para referirse al proceso de obtención de una imagen HDR a partir de una serie de fotografías con diferentes ex-

¹En fotografía, el término de exposición hace referencia a la cantidad total de luz que se permite que alcance el medio fotográfico (película, en fotografía analógica, o sensor, en la digital) durante el proceso de tomar una fotografía. En la Figura 1.2 se observa el efecto de aumentar el tiempo de exposición al tomar fotografías en el mismo escenario.

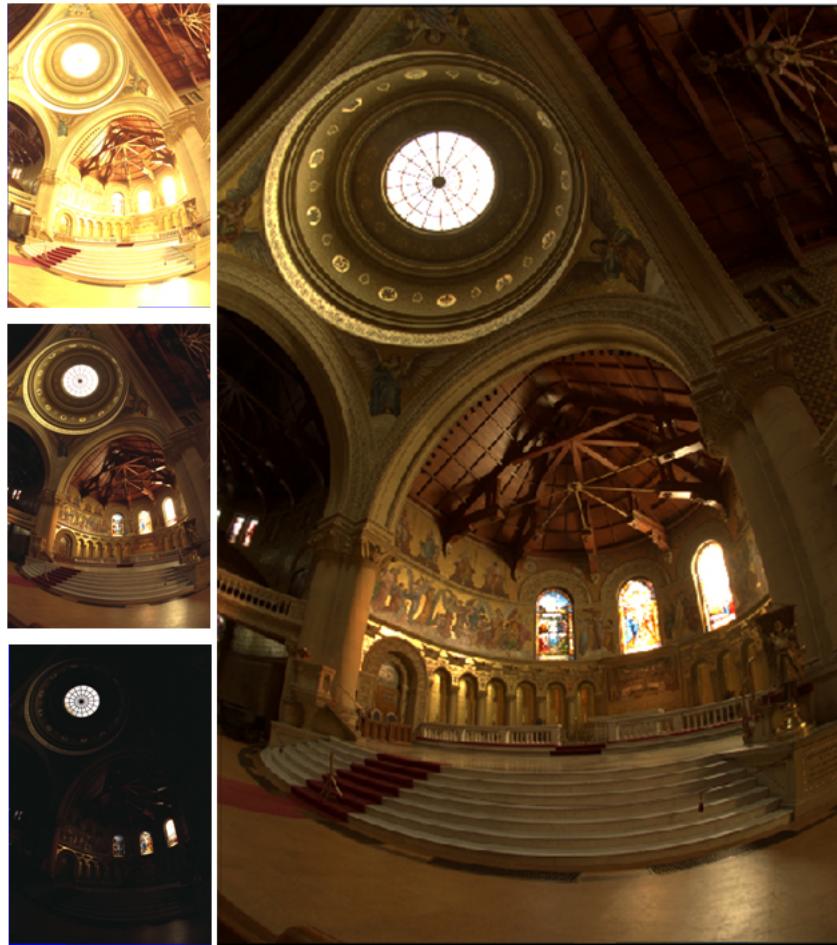


Figura 1.1: Se puede obtener una imagen HDR a partir de capturas de la misma escena con distintos valores de exposición.

posiciones (Figura 1.1). Esta técnica de composición genera resultados más pobres que la producción de una imagen tomada con una cámara con sensor de alto rango dinámico. Sin embargo, y debido al alto precio de estos dispositivos, esta última opción no está al alcance del gran público, a día de hoy.

Además, y tal como se explica de forma detallada en el Apéndice B, los monitores están evolucionando de forma increíble gracias a la tecnología HDR. Los monitores convencionales (Low Dynamic Range o LDR) son capaces de conseguir variaciones lumínicas en ratios de 2000:1. Sin embargo, los nuevos dispositivos de visualización (como el BrightSide DR37-P) pueden alcanzar ratios de contraste que superan 200.000:1 [HDRA], similares al rango visible por el ojo humano [Wan95].

Otro concepto interesante, cuando se habla de HDRI, es el de Mapeado de tono (*Tone Mapping*)[DCWP02a]. En los últimos años, con la creciente necesidad y el aumento de popularidad del HDRI, la tarea de reducir drásticamente el contraste de las imágenes HDR para reproducir rangos lumínicos tan vastos en monitores comunes (dispositivos LDR) se

ha convertido en un tema candente. Este problema, sin embargo, desaparecerá progresivamente cuando los monitores HDR alcancen el mercado de forma masiva [SHS⁺04a].

Más importante si cabe que el *Tone Mapping*, es el *Reverse Tone Mapping*², que surge como respuesta a la pregunta de **qué hacer con todas las imágenes y películas (que están, casi en su totalidad, capturadas en bajo rango dinámico) cuando hayan de ser mostradas en la nueva generación de monitores HDR**. ¿Se puede mostrar todo este contenido legado de forma correcta en los monitores HDR? El éxito de esta tecnología depende de que seamos capaces de encontrar una respuesta adecuada a tal necesidad.

1.2. Objetivos

Existen diferentes *Reverse Tone Mapping Operators* (rTMO's). Se han desarrollado numerosos algoritmos de expansión (lineales [AFR⁺07], exponenciales, basados en mapas de luz [RTS⁺07], operando de forma local -sobre cada píxel de forma independiente-, o de forma global) y estudios comparativos entre ellos; en su mayoría, de carácter psicofísico³.

La mayoría de estudios psicofísicos existentes a día de hoy se dedican a valorar el resultado de aplicar los diversos rTMO's existentes a imágenes correctamente expuestas, debido a que los estudios sobre rTMO's se han desarrollado, hasta el momento, para imágenes *perfectas* (ausencia de ruido, alta resolución, correcta exposición, no compresión). **Hasta la fecha no se conoce ningún estudio que resuelva el problema de cómo expandir el contenido dinámico de imágenes imperfectas** (entendiendo por imperfectas fotografías en baja resolución, comprimidas, con exposición incorrecta, ...).



Figura 1.2: Al aumentar el tiempo de exposición se consiguen fotografías más luminosas.

Este proyecto estudia el problema específico del *Reverse Tone Mapping* para imágenes legadas imperfectas, centrándose en el problema de la incorrecta exposición (ya sea por

²Algunos autores [BLDC06, AFR⁺07] se refieren al proceso como *inverse tone mapping*, mientras que otros [RTS⁺07] usan el término *reverse tone mapping*. Dado que es un campo de reciente aparición no se ha elegido una nomenclatura fija todavía. En esta memoria se opta por utilizar el término *reverse*, dado que el término *inverse* puede referirse a invertir matemáticamente un operador de reproducción de tono, y no al proceso completo de transformación de una fotografía. No obstante, en castellano se opta por utilizar el término *inverso*.

³La psicofísica se aplica a magnitudes que no pueden medirse de forma objetiva. El Apéndice C profundiza en los conceptos psicofísicos.

exceso -sobreexposición- o por defecto -subexposición- de la misma). En primer lugar, se diseñará y realizará un estudio psicofísico que tenga por fin clasificar las imágenes utilizadas como estímulos en correctas, subexpuestas y sobreexpuestas. Acto seguido se mostrarán los resultados de éste estudio junto con algunas estadísticas de primer orden de las imágenes, en un intento de comprender los motivos que ocasionan que una fotografía sea clasificada en uno u otro grupo.

La segunda parte del proyecto estará centrada en el desarrollo de una metodología para evaluar el rendimiento de los rTMO's existentes cuando han de aplicarse a imágenes incorrectamente expuestas. Una vez diseñada esta metodología, se pondrá en práctica el estudio. De él se obtendrán conclusiones interesantes de cara al diseño de un rTMO que sea capaz de producir buenos resultados para cualquier tipo de imagen, atendiendo a la exposición de la misma. (Esto se puede ver, a alto nivel, en la Figura 1.3).

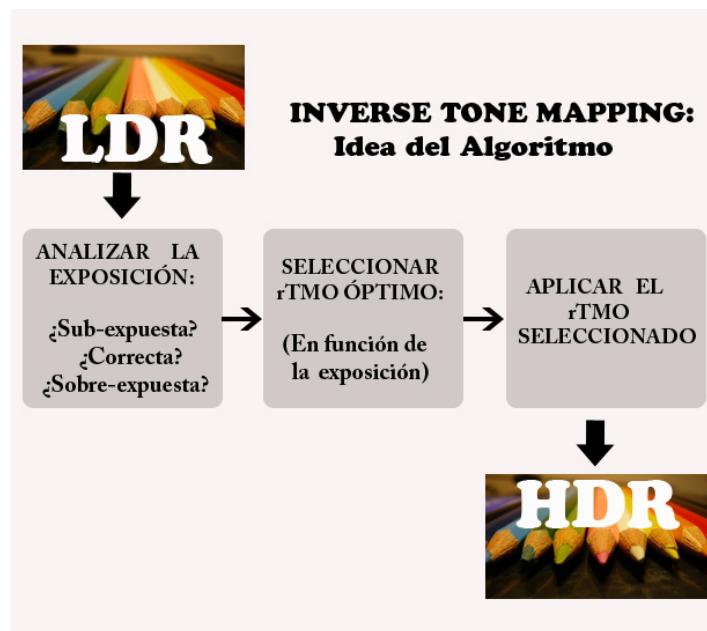


Figura 1.3: Idea del algoritmo-objetivo final.

1.3. Alcance

Para satisfacer los objetivos detallados en el apartado anterior deben superarse varias etapas:

- Toma de fotografías (horquillado de exposición).
- Definición de subexposición y sobreexposición (psicofísica, consulta a especialistas en fotografía).

- Familiarización con hardware HDR, problemas y limitaciones (apoyado por la estancia de una semana en el Max Planck Institut für Informatik. [MPIb]).
- Selección de fotografías que interesan de cara al estudio comparativo entre rTMO's.
- Diseño del primer test psicofísico (para la obtención de una taxonomía de los estímulos).
- Análisis de los resultados del primer test y su relación con estadísticas de primer orden de las imágenes. Establecimiento de hipótesis de causalidad (¿Qué hace que una foto se perciba como correctamente expuesta?).
- Diseño de la metodología y realización del segundo test psicofísico.
- Análisis de resultados y establecimiento de conclusiones.
- Diseño y evaluación del algoritmo final.

Las fotografías tomadas en la primera etapa deben tener una amplia variedad de iluminaciones, cubriendo interiores, exteriores, luz natural (mediodía y atardecer o amanecer) y artificial, objetos cercanos y lejanos, ... Es conveniente resaltar que no se buscan instantáneas ricas en el aspecto artístico, sino que resulten útiles para los tests que se realizarán utilizándolas. Esto incluye todos los grados posibles de sub y sobreexposición (desde extremadamente incorrectas a perfectamente correctas).

También es imprescindible establecer un criterio de qué se considera sub y sobreexpuesto. Para ello se consultará a varios especialistas en el campo de la fotografía y se realizará un test psicofísico con las imágenes anteriores. Este test tiene por objetivo la obtención de una clasificación de las instantáneas en tres grupos: exposición correcta, sub y sobreexpuestas. Cabe destacar que dentro de los dos últimos grupos no todas las imágenes estarán igual de sub/sobreexpuestas. Es decir, algunas serán clasificadas de igual forma por todos los encuestados (el cien por cien de los mismos las englobarán dentro de un mismo grupo), pero ésto no sucederá en la mayoría de los casos. Por tanto, encontraremos fotografías más sub/sobreexpuestas que otras.

Otro paso fundamental para la comprensión y familiarización ha consistido en una estancia de una semana en el Max Planck Institut für Informatik, (Saarbrücken, Alemania). [MPIb] . Esos días, tutorados por Karol Myszkowski [Kar] (especialista en HDR) han sido cruciales para observar las características, limitaciones y problemas de los monitores HDR, aplicar diversos rTMO's, familiarizarse con la problemática existente y realizar las primeras pruebas de visualización de imágenes imperfectas (sub y sobreexpuestas), así como la pre-selección de las imágenes a utilizar en los tests posteriores.

Nótese que no se desea restringir el estudio a las imágenes extremadamente mal expuestas, sino que se pretende ampliar éste a imágenes de todo tipo. Así, se podrán buscar relaciones entre el porcentaje de encuestados que engloban una fotografía dentro de un grupo y el rTMO que mejor la transforma. (Por ejemplo: "Cuando la fotografía se

considera sobreexpuesta por un porcentaje Y de los encuestados, el algoritmo que mejor funciona es el X.”, ver Capítulos 2 y 4).

Otro aspecto fundamental de este proyecto será la preparación de los tests psicofísicos que permitirán establecer relaciones como la comentada. Para ello se habrán de tener en cuenta cuestiones psicológicas (qué preguntar, a quién y cómo hacerlo) y buenas prácticas psicofísicas (aleatorización de las muestras, realización de fase de entrenamiento en el test, etc. Ver Apéndice C).

La ejecución de estos tests será un monitor HDR (BrightSide DR37-P, ver Apéndice B). Nótese que no se dispone de monitores de este tipo en la Universidad de Zaragoza. Esto añade una nueva dificultad, pues los test se diseñarán aquí, pero serán realizados en Alemania, usando el monitor cedido por el Max Max Planck Institut für Biologische Kybernetik [MPIa], bajo la supervisión de Roland Fleming⁴

De cara a la realización de los tests, ha sido de vital importancia la participación en este proyecto de Francesco Banterle y Olga Sorkine, que proporcionaron el resultado de aplicar dos de los rTMO's evaluados (Francesco devolvió las imágenes tras aplicarlos el rTMO que se detalla en [BLDC06] y Olga los resultados de aplicar [RTS+07]). Olga Sorkine⁵ ha colaborado, además, en la elaboración del artículo de investigación presentado al CEIG 2008 [MFSG08] (pendiente de publicación y adjuntado como Anexo E en este mismo proyecto).

Una vez realizados los tests será necesario evaluar estadísticamente los resultados obtenidos, a fin de obtener las conclusiones que nos permitirán conocer qué algoritmo funciona mejor con cada exposición. Con este conocimiento se desarrollará y evaluará un rTMO que obtenga buenos resultados para cualquier exposición fotográfica.

1.4. Estructura de la memoria

1.4.1. Capítulo 2

El problema en detalle: ¿qué hacer con las imágenes legadas (en su mayoría, en formatos de bajo rango dinámico) cuando deban mostrarse en monitores de alto rango?

⁴Roland Fleming ha publicado multitud de artículos en el campo de la percepción [FTA04, FABJ04] y de tests psicofísicos relacionados con HDR [AFR⁺07], por lo que su ayuda y consejos resultan de gran utilidad.

⁵Olga Sorkine ha sido premiada con el *Eurographics Young Researcher Award 2008* y autora de artículos de reconocido prestigio en materia de gráficos por ordenador, y especialmente en modelado geométrico y manipulación de formas e imágenes, como [BS08, ESA07, BSS07]

1.4.2. Capítulo 3

Test psicofísico 1: estableciendo una taxonomía de los estímulos.

1.4.3. Capítulo 4

Test psicofísico 2: examinando el rendimiento de diversos rTMO's

1.4.4. Capítulo 5

Diseñando un nuevo rTMO: partiendo del conocimiento obtenido en anteriores apartados (qué rTMO funciona mejor para cada *tipo* de fotografía y por qué) se plantean las bases para el diseño de un nuevo rTMO.

1.4.5. Capítulo 6

Conclusiones, problemas y trabajos futuros.

1.4.6. Apéndice A

Conceptos fotográficos: conceptos básicos de fotografía (ahondando en aquéllos necesarios para comprender HDR).

1.4.7. Apéndice B

HDR en el pipeline fotográfico. Actualidad y tendencias futuras de captura, almacenamiento, edición y visualización fotográfica.

1.4.8. Apéndice C

Conceptos psicofísicos: conceptos de psicofísica, psicología y percepción.

1.4.9. Apéndice D

Relación de estímulos utilizados, junto a los datos estadísticos más relevantes asociados a los mismos.

1.4.10. Apéndice E

Artículo de investigación presentado al CEIG 2008 (pendiente de aceptación a fecha de cierre de este proyecto).

Capítulo 2

El problema

El primer problema al que se enfrenta este proyecto consiste en establecer una definición de subexposición y sobreexposición, de forma que se pueda establecer un criterio común de cara a las futuras pruebas.

El segundo problema consiste en adecuar la visualización de imágenes legadas sub y sobreexpuestas a los monitores de nueva generación. En primera instancia se estudian las soluciones existentes hoy en día, comparando, para ello, los *Reverse Tone Mapping Operators*, (rTMO's) desarrollados hasta el momento. Una vez descubiertas las bondades y defectos de los distintos algoritmos cuando tratan con contenido imperfecto (limitado, en este estudio, a sub y sobreexposición), el proyecto abordará el desarrollo de un nuevo rTMO.

2.1. Definiendo subexposición y sobreexposición

2.1.1. Introducción

Exposición, para la Real Academia de la Lengua, es la *Acción de exponer a la luz una placa fotográfica o un papel sensible durante cierto tiempo para que se impresione*. Subexposición, por tanto, es la *exposición insuficiente del material fotosensible* y la sobreexposición la *exposición excesiva del material fotosensible* [RAE]. Para algunos casos, esta clasificación es clara (Figura 2.1). Sin embargo, no se ha definido hasta ahora un valor objetivo que determine cuál es la frontera entre la *exposición insuficiente*, la correcta y la *excesiva*.

Además, estos sucesos no son excluyentes entre sí. De hecho, incluso dentro de una misma fotografía pueden existir zonas mal expuestas y otras con exposición correcta (Figura 2.2); y lo que es más, subexposición y sobreexposición en la misma fotografía (Figura 2.3).

Si la corrección de la exposición en una imagen pudiera ser determinada de forma objetiva utilizando únicamente datos de la imagen (sin interpretación humana), el *firmware*



Figura 2.1: La fotografía de la izquierda tiene exposición correcta. Además, este tipo de fotografías se denominan de clave alta, pues la media de los valores lumínicos de los píxeles es alta; es decir, son imágenes en que predominan los tonos claros. En la fotografía de la derecha la exposición es incorrecta. El negativo ha recibido demasiada luz, lo que deriva a pérdida de detalle y contraste. Nótese la diferencia de concepto entre clave alta y sobreexposición.



Figura 2.2: La farola tiene una exposición correcta, sin embargo los edificios del fondo están sobreexpuestos (San Francisco).

de las cámaras podría, en teoría, obtener la exposición óptima para cada escena. La mayoría de las cámaras del mercado ofrecen una estimación que funciona suficientemente bien para la gran mayoría de las escenas e iluminaciones, pero en ocasiones es necesaria la intervención humana para obtener el resultado deseado, especialmente a nivel profesional.

La única forma de que una cámara LDR (con limitación de rango) realizara fotografías perfectas sería que fuera capaz de determinar el motivo de la fotografía, además de in-



Figura 2.3: Subexposición y sobreexposición en la misma fotografía: la zona de la ventana está sobreexpuesta y la del sofá, subexpuesta.

interpretar los datos lumínicos objetivos (distribución del histograma, media, mediana, desviación típica de luminancias, ...) de la escena. A día de hoy resulta impensable que la inteligencia de las cámaras sea tal que sean capaces de reconocer el motivo de la fotografía.

Pero se puede ir más allá. Imagine el siguiente "ejemplo de manual": sea una escena compuesta por un verde paisaje, un coche rojo y un hombre conduciéndolo. Si el fotógrafo desea que el coche esté correctamente expuesto, entonces debe medir la luz reflejada por el coche y subexponer la lectura del fotómetro entre uno y dos *stops*. Sin embargo, si desea que sea el conductor quien esté correctamente expuesto, deberá sobreexponer alrededor de 1,5 *stops*, y si lo que quiere es que sea la hierba la que cuente con una exposición correcta, deberá utilizar la medida exacta del fotómetro. Así, incluso aunque la cámara fuera capaz de interpretar elementos semánticos de alto nivel de la escena, y reconocer patrones (un paisaje verde, un coche rojo, una persona), sería incapaz de adivinar la intención del fotógrafo: ¿en qué elemento quiere éste centrar su atención?

La labor de clasificar una imagen utilizando datos objetivos de la misma es, como se ha expuesto, muy compleja. Resulta interesante apoyar esta argumentación con un ejemplo más. Obsérvense las dos fotografías, tomadas con una cámara convencional, de la Figura 2.4. Ambas cuentan con histogramas de forma muy similar, con medias y medianas de valores lumínicos muy parecidas. Sin embargo, teniendo en cuenta indicadores de alto nivel (reconocimiento del motivo), la fotografía de la izquierda puede ser considerada como correcta (hablando en términos de exposición), mientras que la instantánea de la derecha está claramente sobreexpuesta.

Para poder determinar el grado de corrección en la exposición de una fotografía, son necesarios indicadores semánticos de alto nivel (reconocimiento de patrones, entre otros), además de datos objetivos lumínicos de la escena (forma del histograma, media, mediana,



Figura 2.4: Los histogramas y datos objetivos de luminancias (media, mediana, moda, desviación típica) de los histogramas son muy similares. Sin embargo, la fotografía de la izquierda se puede considerar correctamente expuesta, mientras que la de la derecha está claramente sobreexpuesta.

moda, desviación típica del mismo). Esta afirmación se ve respaldada por los experimentos llevados cabo por Akyüz y otros en [AFR⁺07]. Los autores utilizaron una secuencia LDR con bracketing de exposición para crear las imágenes HDR, tal como se propone en [DM97]. Se pidió a los participantes que determinaran qué exposición de las presentadas para cada escena era la mejor. Los resultados (no incluidos en el artículo, pero disponibles en [Aky]) ponen de evidencia que los encuestados no siempre escogen la imagen con menos pixels sub o sobreexpuestos, ni tampoco la exposición media de la secuencia. Una interpretación de alto nivel (y probablemente, individualizada) parece tener lugar en este proceso de toma de decisiones. El diseño del primer test psicofísico está motivado por estos descubrimientos.

2.1.2. Clasificación de imágenes: tests psicofísicos

La tarea de juzgar si una fotografía es correcta es, como se ha explicado, algo compleja. La única forma de realizar una clasificación *fiable* pasa por realizar un test psicofísico. En estos tests se muestran una serie de fotografías y los encuestados deben clasificarlas, según su criterio, en sub/sobreexpuestas o *correctas*.

Una vez desarrollado el test, se recogen y agrupan los resultados. Así, se dispone de una valoración estadística $(\alpha(F_i), \beta(F_i), \gamma(F_i))$ de cada fotografía F_i , con $0 \leq \alpha(F_i) \leq 1$, $0 \leq \beta(F_i) \leq 1$, $0 \leq \gamma(F_i) \leq 1$ y $\sum (\alpha(F_i), \beta(F_i), \gamma(F_i)) = 1$, siendo $\alpha(F_i)$ el porcentaje de encuestados que clasificaron como subexpuesta, $\beta(F_i)$ el porcentaje de clasificación como correcta, y $\gamma(F_i)$ el de sobreexpuesta.

La metodología utilizada en estos tests se encuentra detallada en el Apéndice C. Los detalles específicos, así como los resultados y conclusiones obtenidas en este test se relatan en el Capítulo 3.

2.2. Diseñando un nuevo rTMO

2.2.1. Comparación de rTMO's existentes

Una vez clasificadas, las imágenes se utilizarán para realizar otro test psicofísico. Los estímulos utilizados en este segundo test serán las imágenes obtenidas tras aplicar distintos rTMOs a cada una de las imágenes del primer test. El medio utilizado será un monitor HDR (BrightSide DR37-P, ver Apéndice B). Nótese que no se dispone de monitores de este tipo en la Universidad de Zaragoza. Esto añade una nueva dificultad (ya comentada en el Capítulo 1), pues los test deberán ser diseñados aquí, pero deberán ser realizados en el monitor cedido por el Max Plankt Institut, que se encuentra en Alemania.

Los encuestados en este nuevo test deberán puntuar cuál de los rTMOs aplicados obtiene *mejores* resultados (más naturales, más agradables a la vista: en definitiva, cuál utilizarían para cada fotografía, por considerar que obtiene resultados óptimos).

Cuando se computen los datos de todos los participantes se contará, para cada imagen, una lista ordenada de preferencia de los distintos rTMO's aplicados. Por ejemplo, para la fotografía F_i , si los operadores preferidos son $rTMO_a, rTMO_c, rTMO_b$ y $rTMO_d$, en este orden, $\sigma(F_i)$ sería.

$$\sigma(F_i) = \{rTMO_a \ rTMO_c \ rTMO_b \ rTMO_d\} \quad (2.1)$$

Los detalles (rTMO's y metodologías utilizadas), resultados y conclusiones relacionadas con este test se enuncian en el Capítulo 4.

2.2.2. Adecuación de cada ITMO en función de la exposición

Del primer test se obtienen conclusiones acerca de la exposición de cada una de las fotografías utilizadas como estímulo. Del segundo test se obtienen conclusiones acerca de la adecuación de cada uno de los rTMO's para cada una de las fotografías (4.3). Se pretende relacionar la información del primer test con la del segundo (por medio de la función $\tau(\alpha, \beta, \gamma)$), de forma que se establezcan conclusiones que relacionen exposición con rTMO preferido, para cada instantánea. De este modo, para una fotografía F_j con los porcentajes $(\alpha_j, \beta_j, \gamma_j)$, se tendría una lista ordenada de preferencias de algoritmos, al estilo de

$$\tau(\alpha_j, \beta_j, \gamma_j) = \{rTMO_a \ rTMO_b \ rTMO_c \ rTMO_d\} \quad (2.2)$$

2.2.3. Diseño del algoritmo

De los anteriores tests se tiene una base de conocimiento (evidenciada por la función τ) para las imágenes utilizadas como estímulo. Por tanto, para una nueva fotografía F_k , el proceso de obtención del mejor algoritmo pasaría por:

- Obtener los porcentajes $\alpha(F_k)$, $\beta(F_k)$, $\gamma(F_k)$ (por psicofísica o utilizando indicadores semánticos de alto nivel en combinación con datos objetivos de la imagen, como se apuntaba al principio de este Capítulo).
- Utilizar los coeficientes anteriores como entrada de la función τ .
- Aplicar el primer método de lista de rTMO's que devuelve la función τ .

Capítulo 3

Clasificando las imágenes: test psicofísico 1

El principal propósito de este test psicofísico es obtener una taxonomía de los estímulos empleados, de tal forma que éstos se clasifiquen en subexpuestos, correctos o sobreexpuestos, cuando la mayoría de los encuestados así lo hagan. Además se aprovecha la ocasión para verificar algunas de las hipótesis presentadas en otros estudios (por ejemplo, los de Akyüz y colaboradores en [AFR⁺07]) y obtener algunas conclusiones de causalidad, del tipo *¿qué hace que una fotografía se englobe en un grupo, y no en otro?*.

En este Capítulo se detallan los estímulos utilizados, la metodología empleada, así como los resultados y conclusiones obtenidas.

3.1. Diseño del test

3.1.1. Motivación

Tal como se ha discutido en el primer apartado del Capítulo 2, la sub y sobreexposición no han sido definidas de forma objetiva. Esto sugiere que no existe correlación entre la aparente corrección de exposición y valores objetivos de datos de la imagen, como histograma de luminancias, media, mediana o porcentaje de pixels que no alcanzan (pixels subexpuestos) o superan (pixels sobreexpuestos) determinado valor de luminancia.

En teoría sería posible detectar un subconjunto de casos en que la imagen es correcta, por ejemplo cuando el histograma muestra valores nulos por debajo y por encima de ciertos umbrales. Sin embargo, aplicando este criterio se obtendrían falsos positivos, como por ejemplo en el caso de imágenes de bajo contraste con superficies uniformemente iluminadas. Para algunas aplicaciones, una aproximación útil podría consistir en definir un umbral inferior, por debajo del cual los pixels se consideran subexpuestos, y otro superior, a partir del cual los pixels se consideran sobreexpuestos (tal y como se hacen Meyland y otros en [MDS06]). Sin embargo estas operaciones se realizan a nivel de pixels y no



Figura 3.1: Estímulos: un representante de cada escena.

proveen de información a más alto nivel, como por ejemplo de apariencia global de la fotografía. Además, y tal y como se comenta en el Capítulo 2, en una misma foto pueden existir áreas subexpuestas y otras sobreexpuestas.

Por todo ello parece que para clasificar una imagen correctamente en uno de los grupos se necesitan observar semánticas de alto nivel, dependientes del contexto, como se sugiere en estudios previos ([AFR⁺07, Aky]). Se va a someter esta hipótesis a examen, comparando la percepción subjetiva de la exposición en imágenes con estadísticas de primer orden de las mismas: histograma de luminancias, media, mediana y porcentaje de pixels sobreexpuestos y subexpuestos (definidos a partir de ciertos umbrales).

Nótese, sin embargo, que la principal motivación para la realización de este test no es establecer hipótesis de causalidad ("Cuando la fotografía se percibe como X es porque el dato objetivo Y verifica la condición Z"), sino la obtención de una clasificación de los estímulos, de cara a ser utilizada en la segunda parte de este proyecto.

3.1.2. Estímulos

Se usarán imágenes tomadas en diez escenarios distintos. Estas imágenes fueron capturadas por una Nikon D200 a resolución 3872 por 2592 y re-escaladas a 1920 por 1080 para propósitos de visualización. Las fotografías se eligieron de forma que se cubriera un amplio abanico de iluminaciones y medios. Se realizó bracketing de exposición de cinco disparos por secuencia, desde claramente subexpuesto (etiquetadas como 1) a altamente sobreexpuesto (etiquetadas como 5), obteniéndose un total de 50 imágenes.

En la Figura 4.7 se observa una fotografía representante de cada una de las secuencias elegidas. La Figura 4.8 muestra una secuencia completa (las cinco exposiciones distintas de una misma escena).

Los valores de luminancia se obtienen, tal como se propuso en [RWPD05], según la



Figura 3.2: La secuencia completa de bracketing para la escena *sunset*, junto a sus respectivos histogramas de luminancias.

siguiente fórmula:

$$L(Pixel_{ij}) = 0.213 * R + 0.715 * G + 0.072 * B \quad (3.1)$$

con R el valor numérico de rojo, G el de verde y B el de azul, para el $Pixel_{ij}$.

Las Tablas 3.1 y 3.2 presentan las medias y medianas de luminancia, respectivamente. La Tabla 3.3 muestra el porcentaje de pixels con luminancia mayor o igual a 254. La elección de este umbral se fundamenta en el buen funcionamiento del mismo, cuando se trata de discriminar áreas sobreexpuestas en fotografías, tal como se propone en [RTS⁺07]. Finalmente, la Tabla 3.4 muestra el porcentaje de pixels con luminancia nula, lo que representa el umbral elegido de subexposición. Otras estadísticas de las imágenes se pueden consultar en el Anexo D.

Secuencia / Exposición	1	2	3	4	5
Building	126.66	158.96	183.77	211.39	232.45
Car	14.72	23.79	38.41	58.58	85.06
Indoor flower	19.67	30.29	50.28	60.41	60.51
Lake	70.77	86.47	138.04	170.17	201.54
Pencils	26.68	43.37	68.65	104.5	138.95
Computers	14.93	20.55	24.05	34.98	52.51
Waxes	29.48	50.41	78.21	111.85	148.84
Sunset	91.86	108.56	139.87	194.62	229.99
Graffiti	127.06	169.69	202.56	226.57	242.94
Strawberries	90.39	130.18	168.60	199.79	223.45

Tabla 3.1: Luminancia media de los píxeles para cada una de las fotografías de cada secuencia.

3.1.3. Metodología

El diseño del experimento psicofísico sigue el esquema conocido como el *método del estímulo constante* [BCJC]: las cincuenta imágenes se muestran una por una, en orden aleatorio, mezclando tanto exposiciones como escenas. Se pide a los participantes que

Secuencia / Exposición	1	2	3	4	5
Building	85	137	189	240	255
Car	6	14	27	50	86
Indoor flower	7	15	33	45	45
Lake	43	61	119	164	210
Pencils	9	18	39	83	136
Computers	0	3	5	11	26
Waxes	23	46	80	120	167
Sunset	75	97	136	218	255
Graffiti	133	189	233	255	255
Strawberries	90	140	198	246	254

Tabla 3.2: Mediana de luminancia de los píxeles para cada una de las fotografías de cada secuencia.

Secuencia / Exposición	1	2	3	4	5
Building	4.60	19.97	40.03	44.33	47.78
Car	0.02	0.03	0.21	0.60	1.90
Indoor flower	0.64	0.79	1.30	1.95	1.98
Lake	0	0	18.71	23.93	34.49
Pencils	0	0	0	0	1.94
Computers	0.14	0.54	0.90	2.32	7.36
Waxes	0	0	0	0.01	2.26
Sunset	0.01	4.45	8.35	26.16	51.72
Graffiti	0.01	1.00	20.19	49.20	61.46
Strawberries	0	0.01	5.97	22.16	38.39

Tabla 3.3: Porcentaje de pixels con valor lumínico 254 y 255.

clasifiquen cada imagen en uno de los siguientes grupos: (1) subexpuesta, (2) correcta, (3) sobreexpuesta. Este test sigue las buenas prácticas psicofísicas enunciadas en el Apéndice C. El encuestado puede visualizar cada una de las fotos tanto tiempo como desee, pudiendo además navegar adelante o atrás entre las imágenes, revisar y cambiar valoraciones de las mismas. Al inicio del test se proporciona la definición de la RAE para exposición, sub y sobreexposición. Además, para asegurar la validez de los datos, los participantes realizan un pequeño test de prueba, como se sugiere en [Ken75]. En él, son invitados a juzgar unas cuantas imágenes hasta que adquieran una cierta destreza tanto en el aprendizaje de los conceptos como en el funcionamiento del test. Estas imágenes previas vienen de escenas extra que no son parte del test en sí mismo.

Para la visualización del test se utiliza un Benq FP241VW de 24 pulgadas. El experimento se lleva a cabo en una habitación oscura para no reducir el ratio de contraste del monitor (alrededor de 60:1). La luminancia ambiental, medida en la pared de la sala, es de 26 cd/m^2 . La distancia del participante al sistema de visualización será, aproximadamente, 50 cm. El interfaz de usuario se muestra en la Figura 3.3.

Secuencia / Exposición	1	2	3	4	5
Building	0	0	0	0	0
Car	30.66	19.65	9.23	4.08	1.16
Indoor flower	26.11	17.10	5.99	3.62	3.50
Lake	0.10	0	0	0	0
Computers	58.06	23.53	14.11	2.70	0.12
Waxes	13.13	4.56	0.45	0	0
Sunset	0	0	0	0	0
Pencils	12.51	8.38	5.04	1.02	0.06
Graffitti	0	0	0	0	0
Strawberries	0	0	0	0	0

Tabla 3.4: Porcentaje de pixels con valor lumínico nulo.

Los encuestados son un grupo de 24 hombres y mujeres, todos con visión normal o corregida a normal. La mitad de ellos se sienten aficionados y con conocimientos básicos de fotografía, y la otra mitad indiferentes.

3.2. Resultados y conclusiones

La Tabla 3.4 muestra el resultado numérico de los tests. La misma información puede observarse también de modo gráfico en la Figura 3.5.

El análisis de resultados se hará en cuatro fases. En la primera, se analiza la forma del histograma en relación a los resultados del test. En los siguientes se realiza una comparación similar, alterando el dato objetivo: en la segunda se relaciona con la media, en la tercera con la mediana y en la última con el porcentaje de pixels sobreexpuestos y subexpuestos.

La inspección visual de los resultados del test (Figura) muestra, como era de esperar, la distribución diagonal lógica de la exposición percibida. La fuerte ausencia de luz sobre algunos objetos principales de la escena ha sido interpretada mayoritariamente como subexposición (*indoor, car*), mientras que iluminaciones muy altas se perciben como sobreexposición (*building, sunset*), aunque podría ser que fuera la intención del fotógrafo conseguir dichos efectos. Esto vuelve a poner de manifiesto la necesidad de semánticas de alto nivel y posiblemente intervención humana para juzgar la corrección de la exposición. Algún tipo de sistema inteligente de aprendizaje, como los *Vector Machines* [Vap95], podría resultar adecuado para separar de forma óptima las imágenes percibidas como sub o sobreexpuestas, o incluso en correctas e incorrectas (tanto sub, como sobreexpuestas). Cuatro de las cinco imágenes con gradientes más fuertes (las cuatro mencionadas anteriormente y *computers*) obtuvieron el menor número de votos como *correcta*, sugiriendo que estadísticas de segundo orden podrían brindar algo de información adicional en este sentido. Como se esperaba, el histograma por sí mismo no da suficiente información sobre

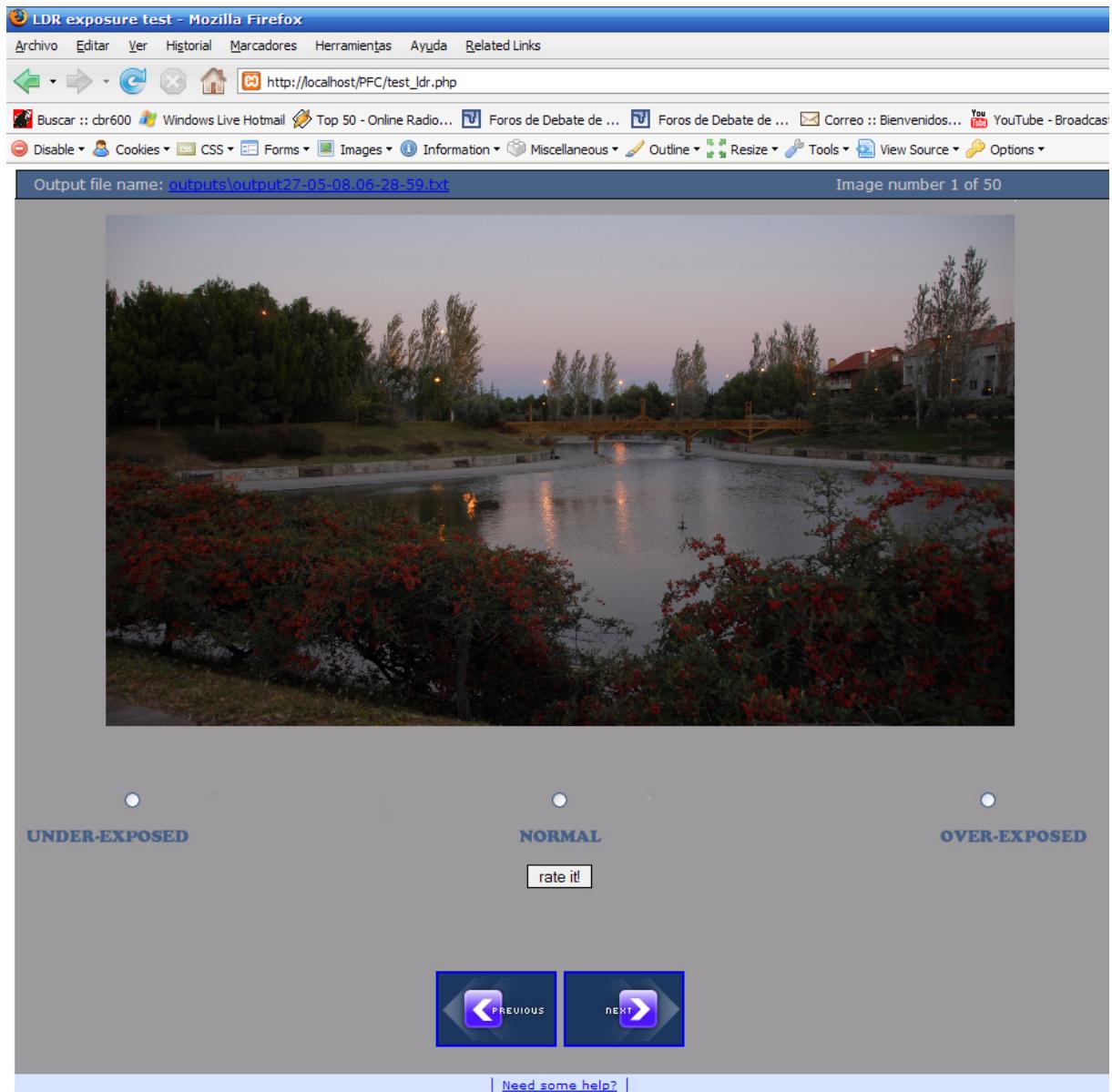


Figura 3.3: Interfaz de usuario utilizado en el primer test psicofísico. La aplicación web se ha programado siguiendo las normas de usabilidad y accesibilidad web y pasó con éxito el test de la aplicación TAW [Web] para el nivel de accesibilidad AA.

Scene name		Exp. 1	Exp. 2	Exp. 3	Exp. 4	Exp. 5
building	under	17%	4%	4%	0%	0%
	correct	50%	30%	17%	9%	0%
	over	33%	66%	79%	91%	100%
car	under	100%	96%	83%	83%	46%
	correct	0%	4%	13%	8%	41%
	over	0%	0%	4%	9%	13%
indoor flower	under	92%	95%	42%	25%	4%
	correct	8%	5%	50%	54%	75%
	over	0%	0%	8%	21%	21%
lake	under	42%	13%	0%	0%	0%
	correct	58%	83%	75%	50%	0%
	over	0%	4%	25%	50%	100%
computers	under	58%	71%	25%	7%	0%
	correct	42%	25%	67%	83%	83%
	over	0%	4%	8%	8%	17%
waxes	under	29%	25%	17%	4%	0%
	correct	71%	75%	83%	91%	67%
	over	0%	0%	0%	4%	33%
sunset	under	38%	33%	8%	0%	0%
	correct	29%	42%	17%	5%	0%
	over	43%	25%	75%	95%	100%
pencils	under	75%	46%	21%	0%	0%
	correct	25%	54%	63%	71%	38%
	over	0%	0%	16%	29%	62%
graffitti	under	8%	0%	0%	0%	0%
	correct	92%	92%	46%	21%	0%
	over	0%	8%	54%	79%	100%
strawberries	under	9%	0%	0%	0%	0%
	correct	91%	91%	79%	21%	0%
	over	0%	9%	21%	79%	100%

Figura 3.4: Resultados numéricos del test psicofísico: percepción de la exposición para cada una de las distintas exposiciones de cada una de las secuencias.

la exposición de una imagen.

Existe una correlación positiva entre la media de luminancias de píxeles y la percepción de sobreexposición de la fotografía, más en concreto se ha analizado el coeficiente de correlación de Pearson¹ es $\rho_m^o = 0.869$. De acuerdo a la valoración de Cohen, este coeficiente se considera alto en el campo de la investigación psicológica [Coh88]. Asimismo, existe una correlación similar entre la mediana y la clasificación como sobreexposición $\rho_{md}^o = 0.846$). Además, y como es de esperar, también existe correlación entre la subexposición percibida y los indicadores estadísticos analizados, aunque sorprendentemente, no es tan alta ($\rho_m^u = -0.726$ and $\rho_{md}^u = -0.691$).

Se observa un comportamiento análogo para el porcentaje de pixels incorrectamente expuestos (por superar el umbral de sobreexposición, o por no alcanzar el de subexposición). Existe una fuerte correlación entre la percepción de la sobreexposición y los pixels sobreexpuestos ($\rho_p^o = 0.890$). Tal como sucedía para la media de luminancias, la correlación entre la clasificación como subexposición existe, pero es inferior ($\rho_p^u = 0.675$). Este efecto puede ser debido a la conocida asimetría del sistema visual humano [Liv02].

La Figura 3.6 muestra estos resultados para el caso de la media y subexposición. La Figura 3.7 muestra la relación entre la percepción de la sobreexposición y el porcentaje de pixels con luminancia mayor o igual a 254. Estos dos casos son los que mayores coeficientes de Pearson presentan.

Además, el lector podría pensar que la percepción de la correcta exposición podría estar relacionada con la ausencia de pixels incorrectamente expuestos en la imagen. Existen indicios de esto, tal como indica el coeficiente de correlación entre las dos variables ($\rho_{sum}^c = -0.676$). Sin embargo este coeficiente no se considera alto en el área de investigación psicológica [Coh88], por lo que no pueden extraerse resultados concluyentes.

3.3. Conclusiones

Pese a que la motivación principal de este test ha sido obtener una taxonomía de los estímulos (de cara a la realización de otros tests), se han establecido y sometido a examen hipótesis muy interesantes desde el punto de vista de la comprensión de la percepción de la exposición. La verificación de estas teorías es fundamental de cara al desarrollo de nuevos rTMO's.

En definitiva, las dos ideas principales que se derivan son:

- Los resultados parecen confirmar la hipótesis de que se necesitan semánticas de alto nivel para una clasificación correcta de la exposición. Esta conclusión es muy interesante, pues parece entrar en conflicto con la teoría de que la percepción visual está

¹El coeficiente de correlación de Pearson se define como $\rho_{X,Y} = \frac{E(XY) - E(X)E(Y)}{\sqrt{E(X^2) - E^2(X)} \sqrt{E(Y^2) - E^2(Y)}}$, donde E es la esperanza de una variable. En este caso, X es la clasificación obtenida por psicofísica e Y el dato objetivo con que se está comparando, esto es: media, mediana o porcentaje de pixels sub y sobreexpuestos.

basada en atributos de bajo nivel de las imágenes [AFR⁺07]. Podría argumentarse, sin embargo, que para algunos casos extremos esta afirmación es falsa. Por ejemplo, una fotografía velada² sería clasificada -casi con total certeza- como sobreexpuesta incluso cuando no pudiera reconocerse el objeto fotografiado (y probablemente *debido a* esta ausencia de objetos reconocibles). Sin embargo, la conclusión obtenida es válida para la gran mayoría de las instantáneas, obviando casos extremos como el presentado.

- Se ha observado una notable asimetría entre la percepción de la sobreexposición y la de la subexposición, cuyo origen puede residir en el funcionamiento del sistema visual humano. Es posible que los umbrales elegidos condicionen los resultados.

Como se ha comentado, ambas conclusiones necesitan verse apoyadas por la ejecución de más estudios y pruebas. En este sentido, se considera que existe un amplio potencial de investigación futuro en este área. Asimismo, y como podrá observar el lector, las variables objetivas analizadas (histograma, media, mediana y porcentajes de pixels) están claramente correladas entre sí. Deberían realizarse estudios adicionales para de-correlatar estos parámetros. Deberían analizarse también indicadores estadísticos de alto nivel, dado que la inspección visual sugiere una correlación directa entre la percepción de la exposición y la magnitud de gradientes lumínicos. No obstante, se insiste en que la motivación principal de este test es establecer una clasificación de los estímulos, por lo que los estudios sugeridos quedan fuera del alcance de este proyecto, pudiendo entenderse como una ampliación y continuación del mismo.

²En fotografía tradicional, una instantánea se considera velada cuando ha recibido luz durante el proceso de revelado del negativo

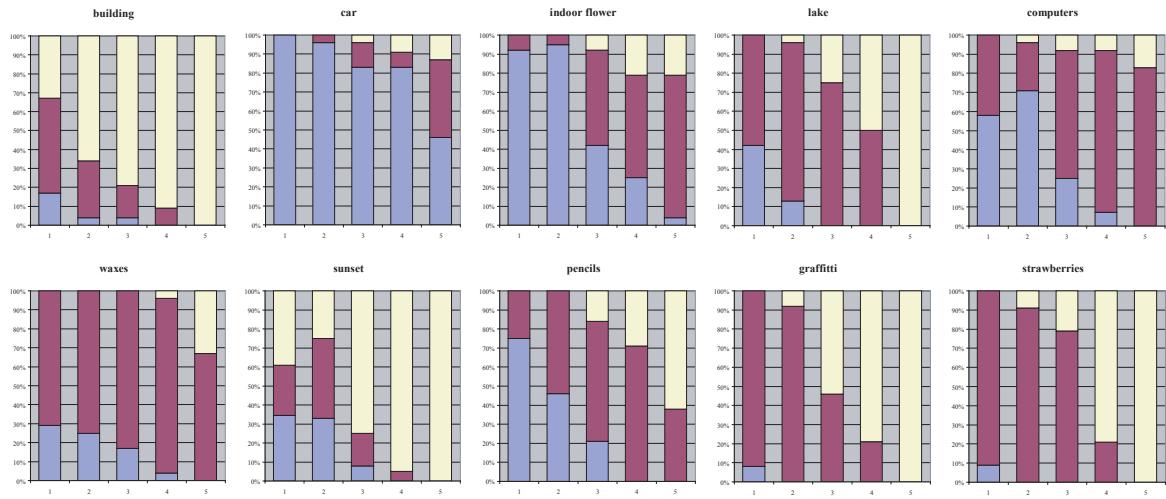


Figura 3.5: Resultados del test psicofísico: taxonomía de los estímulos utilizados. El eje X representa las cinco exposiciones para cada escena, mientras que el eje Y representa el porcentaje acuerdo en la clasificación de los estímulos (morado para la subexposición, amarillo para la sobreexposición y rojo para exposición correcta).

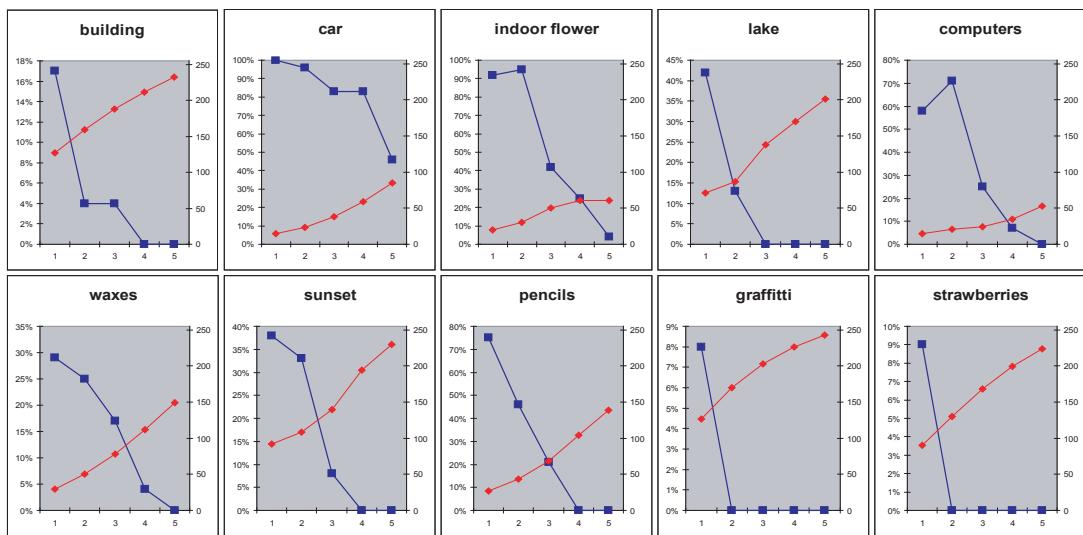


Figura 3.6: Correlación entre los resultados del test psicofísico para subexposición (azul) y la media de luminancia de los pixels de las imágenes (rojo). El eje X representa las cinco exposiciones de cada escena. El eje Y representa el porcentaje de encuestados que perciben el estímulo como subexpuesto (izquierda) y los valores medios de luminancia de los mismos (derecha). Nótese el cambio de escala en el eje Y.

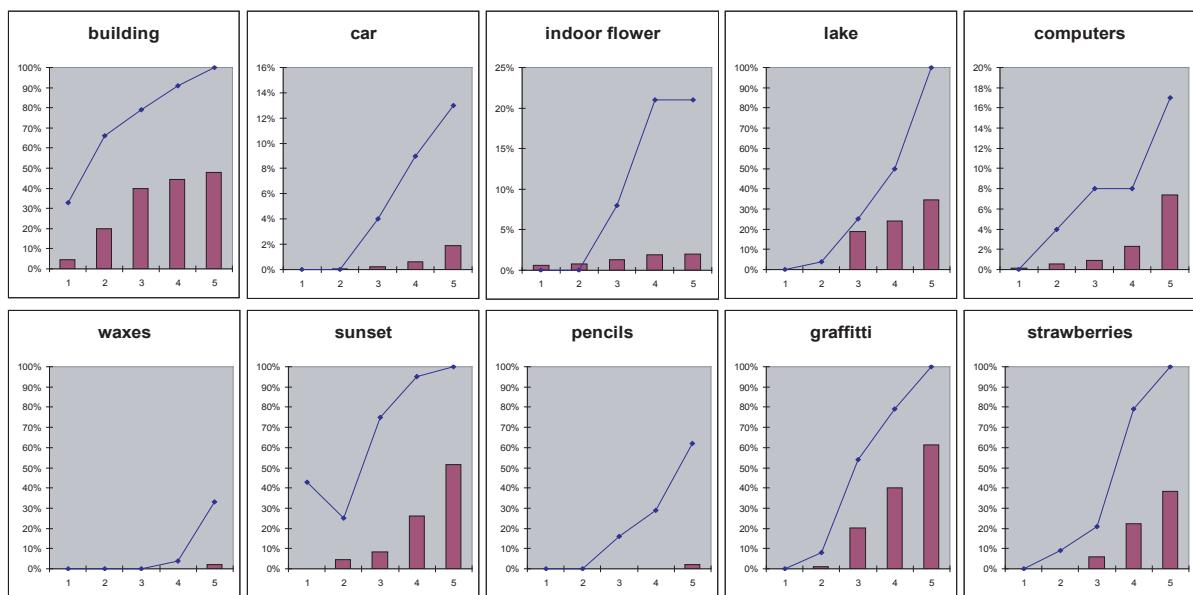


Figura 3.7: Correlación entre los resultados del test psicofísico para sobreexposición (líneas azules) y el porcentaje de pixels sobreexpuestos (barras). El eje X representa las cinco exposiciones de cada escena. El eje Y respresenta el porcentaje de encuestados que perciben el estímulo como sobreexpuesto. Nótese el cambio de escala en el eje Y para las diferentes secuencias.

Capítulo 4

Comparando rTMO's: test psicofísico 2

La facilidad de uso y el abaratamiento de las cámaras y dispositivos de visualización digitales han facilitado el rápido crecimiento de este sector. Los avances en la tecnología posibilitan importantes mejoras en la obtención de nuevas fotografías. Sin embargo, los consumidores, la industria de la fotografía y el cine disponen de gran cantidad de imágenes almacenadas, de las que a menudo es imposible obtener una nueva captura con el hardware que acaba de aparecer en el mercado (La fotografía cuenta con casi 180 años de historia, por lo que muchas instantáneas se han realizado a personajes o lugares que ahora ni siquiera existen). Por tanto, las fotografías legadas deben ser capaces de beneficiarse de las ventajas de los nuevos sistemas de visualización. En la historia reciente encontramos multitud de ejemplos de este fenómeno, siendo algunos de los más notables la *colorización de imágenes en blanco y negro* [WAM02, LLW04] o la *reparación de zonas dañadas de la imagen y eliminación de objetos indeseados en la fotografía* [HT96, BLLC02, BVSO03].

El objetivo de este Capítulo es obtener las claves para aprovechar las ventajas de los dispositivos de visualización HDR con imágenes legadas de contenido no-óptimo; en este caso fotografías subexpuestas y sobreexpuestas. Este proceso se conoce como *reverse tone mapping*, o mapeado de tono inverso.

Los resultados del experimento anterior (Capítulo 3) proporcionan una clasificación de las imágenes en (1) subexpuestas, (2) correctamente expuestas y (3) sobreexpuestas. Una vez realizada esta separación, surgen multitud de preguntas: ¿qué tal se adaptan los *reverse tone mapping operators* (rTMO's) existentes a las instantáneas incorrectamente expuestas? ¿Son capaces de producir resultados aceptables? ¿Bajo qué condiciones funciona bien cada uno de ellos? ¿Hay alguno que supera claramente al resto, bien sea por prestaciones puras o por compromiso entre prestaciones y rapidez en la producción de resultados? El propósito de este Capítulo es dar respuesta a todas estas preguntas.

4.1. Estado del arte: principales rTMO's

4.1.1. *Linear expansion*

La expansión lineal (del rango dinámico) es la forma más simple de obtener una imagen HDR a partir de una única imagen LDR. La Figura 4.1 muestra de forma gráfica este proceso, definido matemáticamente como sigue:

$$L' = k * \left(\frac{L - L_{min}}{L_{max} - L_{min}} \right)^\gamma \quad (4.1)$$

con L la luminancia del pixel a escalar, L_{min} y L_{max} los valores de luminancia mínimo y máximo de la fotografía, k la mayor intensidad de entrada admitida por el monitor HDR. En [AFR⁺07] se realizan comparativas del rendimiento de este operador para distintos valores de γ ($\gamma = 0.45$, $\gamma = 1$, $\gamma = 2.2$), y se concluye que el rendimiento óptimo se consigue para $\gamma = 1$. En el caso que nos ocupa, por tanto, se implementará el rTMO usando este valor del parámetro γ .

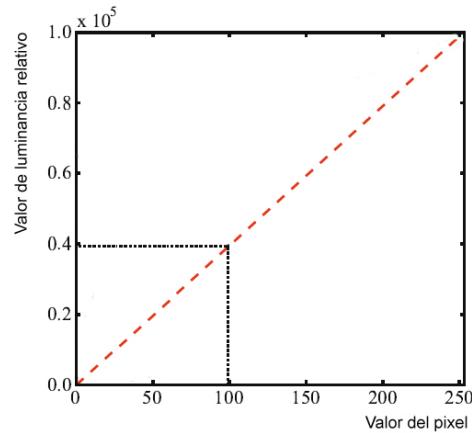


Figura 4.1: Proceso de expansión lineal.

En [AFR⁺07] se afirma que el rendimiento de este rTMO es sorprendentemente bueno, al menos para el conjunto de fotografías testeadas. Otra característica importante de este rTMO es la rapidez en la obtención de la imagen resultado, motivada por la simplicidad del mismo.

4.1.2. *Inverse Tone Mapping algorithm*[BLDC06], por *University of Bristol*

Este rTMO está basado en invertir el operador de reproducción de tono de Reinhard [RSSF02] y combinar su salida con un mapa de densidad de las fuentes lumínicas esti-

madas de la imagen de entrada. Este mapa de luces (originalmente, *expand map*) permite un mayor incremento del rango dinámico, puesto que si únicamente se realizará una inversión del operador de reproducción de tono sólo podría alcanzarse un rango moderado sin causar ruidos en la imagen. Aunque esta aproximación funciona relativamente bien con imágenes estáticas, el *expand map* ocasiona parpadeos en las imágenes, limitando su utilización en secuencias de video.

Los pasos que sigue el método son:

- Generación de una imagen HDR inicial invirtiendo el operador de tono de Reinhard et al [RSSF02].
- Aplicación de *Median Cut Algorithm* [Deb05] para buscar áreas de alta luminosidad.
- Creación de *expand map* a partir de los resultados del algoritmo del paso anterior (ver Figura 4.2).
- Interpolación lineal entre el LDRI con el HDRI del primer paso, utilizando los valores del *expand map* como pesos de interpolación.

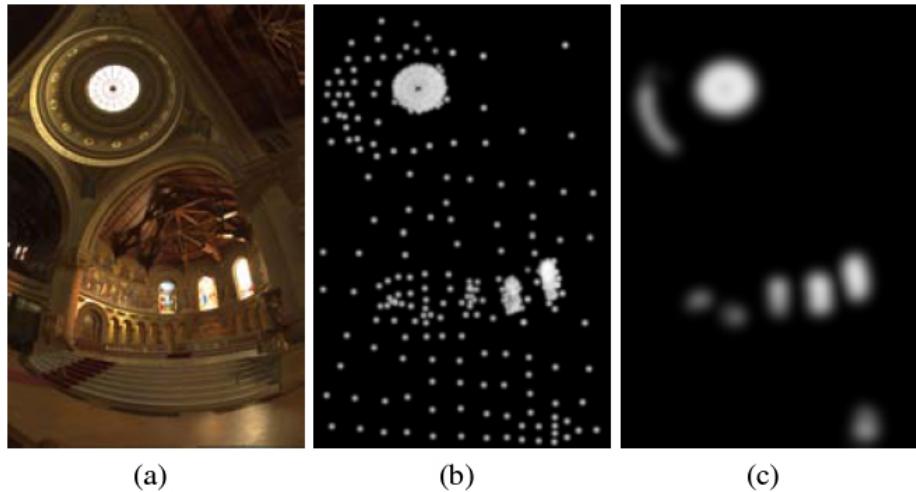


Figura 4.2: Creación del *expand map*: (a) muestra la fotografía LDR de entrada. (b) los resultados de aplicación de *Median Cut Algorithm* [Deb05] para encontrar las fuentes de luz, y (c) el *expand map* que se utilizará para producir la imagen HDR final. Las imágenes fueron extraídas de [BLDC06].

Este método produce resultados aceptables para imágenes estáticas, pero su lentitud en producir resultados, así como la necesidad de estimar parámetros que varían de una fotografía a otra (sólo para la creación del *expand map* hay que ajustar radio y número de fuentes de luz, como mínimo), imposibilitan su uso en aplicaciones en tiempo real. Además, este rTMO opera realizando una segmentación entre distintas regiones de la imagen. Esto se traduce en presencia de ruido y parpadeos entre un fotograma y el siguiente. Otro inconveniente importante de este rTMO, indicado por los autores en el momento de su

publicación([BLDC06]), y sin solución conocida en la actualidad, es que el número de regiones oscuras, normales y claras de la fotografía de entrada debe ser balanceado para que el algoritmo produzca buenos resultados.

4.1.3. *The reproduction of Specular Highlights on High Dynamic Range Displays*[MDS06], por *Ecole Polytechnique Fédérale de Lausanne, Switzerland* y *Sharp Laboratories of America, Camas WA, USA*

Este reproductor de tono se centra en las zonas brillantes de la imagen (*specular highlights* o simplemente, *highlights*) y varía el rango dinámico asignado a ellas de forma distinta a como lo varía para el resto de la fotografía (conocido como *contenido difuso*). Para cada imagen se elige un pixel (de forma manual o utilizando los métodos descritos en [MDS07]), cuya intensidad lumínica determina el umbral entre las regiones difusas y los brillos especulares. Estas dos regiones se escalan linealmente utilizando funciones con distinta pendiente, tal como muestra la Figura 4.3. Este algoritmo encuentra su sentido en las teorías de la percepción de la luz [GKB⁺99]. En ellas, se requiere fijar una correspondencia entre los valores de luminancia de la imagen y los valores lumínicos físicos que representan.

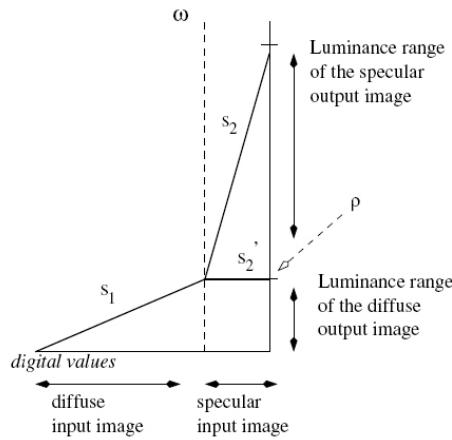


Figura 4.3: El rTMO de Lausanne aplica un escalado lineal con dos pendientes diferentes. Una, menos pronunciada, para las luces difusas. Otra, con más pendiente, para los brillos especulares. El umbral entre una zona y otra (ω) se elige de forma manual. El parámetro ρ simboliza el valor del máximo porcentaje de intensidad lumínica del dispositivo de visualización asignado al pixel con luminancia igual a ω .

La función de escalado de tono se define de forma matemática como

$$f(\delta(p)) = \begin{cases} s_1 * \delta(p) & \text{si } \delta(p) \leq \omega \\ s_1 * \omega + s_2 * (\delta(p) - \omega) & \text{si } \delta(p) > \omega \end{cases} \quad (4.2)$$

, con $s_1 = \frac{\rho}{\omega}$, $s_2 = \frac{1-\rho}{\delta_{max}-\omega}$ y $\delta(p)$ el valor de luminancia del pixel p

La principal desventaja de este algoritmo es la necesidad de fijar el parámetro ρ , cuyo valor idóneo varía para cada imagen. Cuanta mayor es el rango dinámico asignado a la imagen difusa (pixels con luminancia por debajo de ω), más clara está la imagen, pero se decremente el rango asignado a brillos especulares. Esto se muestra en la Figura 4.4, extraída de [MDS06].

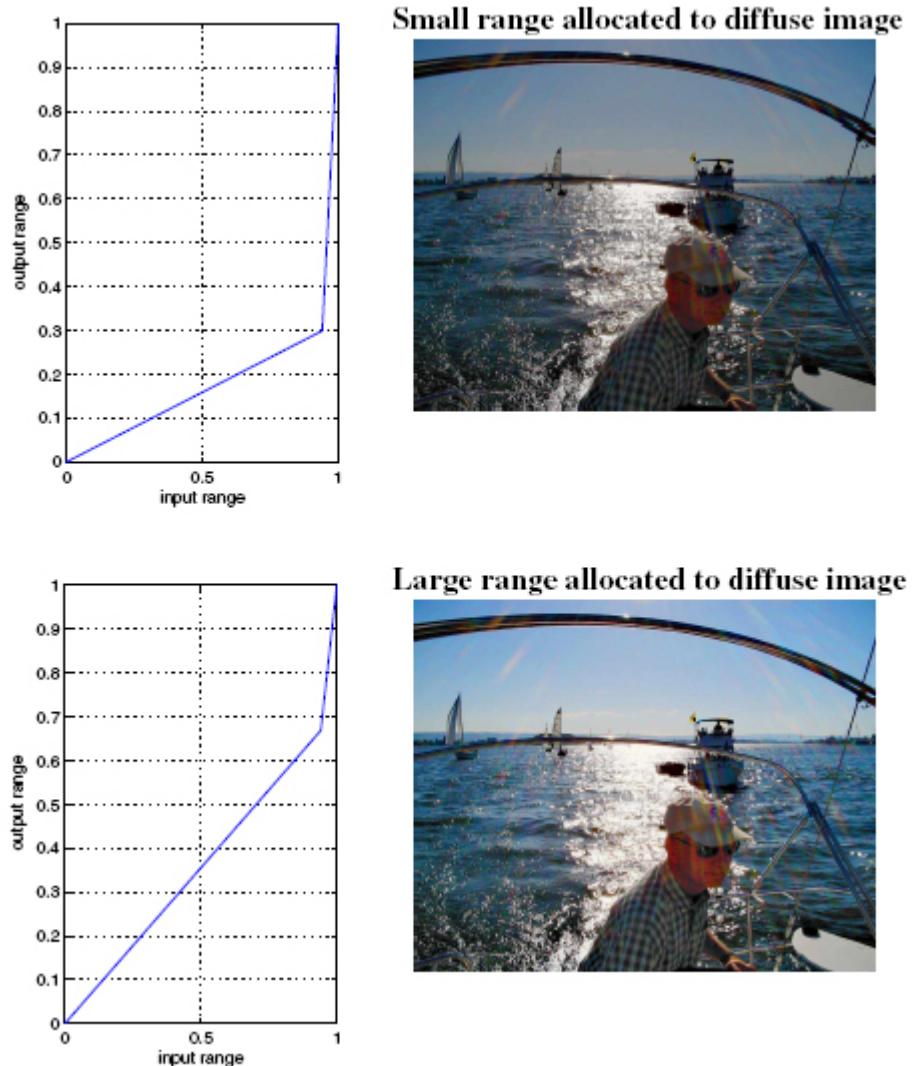


Figura 4.4: Variando el parámetro ρ se obtienen resultados muy distintos. A mayor valor de ρ , más clara es la imagen en conjunto, y menos diferencia hay entre la intensidad de las zonas conocidas como *imagen difusa* y *brillos especulares*. Así, la imagen superior es más oscura que la inferior. El valor del parámetro óptimo varía de una fotografía a otra y debe ajustarse de forma manual.

Este algoritmo obtiene mejores resultados para imágenes oscuras que el escalado lin-

real simple, aunque para imágenes muy iluminadas, el rendimiento de ambos es similar [AFR⁺07].

4.1.4. *Ldr2hdr: On-the-fly Reverse Tone Mapping* [RTS⁺07], por la University of British Columbia y Dolby Canada

Este algoritmo fue desarrollado específicamente para trabajar en tiempo real; esto es, orientado al procesamiento de secuencias de video "al vuelo", sin necesidad de realizar operaciones de pre-computación. Este rTMO es suficientemente rápido como para trabajar en tiempo real con resoluciones de vídeo de calidad HDTV. Además no se precisa el ajuste manual de parámetros variables para cada imagen, ni otro tipo de interacción por parte del usuario (los parámetros de entrada del algoritmo son la imagen a convertir y una serie de datos dependientes del dispositivo de visualización, como la resolución o amplitud del rango dinámico del monitor en cuestión). Es importante notar que este algoritmo no produce diferencias significativas en los colores e intensidades de los fotogramas, salvo que se produzcan en estos mismos, proporcionando secuencias de imágenes coherentes y evitando parpadeos incómodos en la reproducción de secuencias de video. Además los métodos propuestos pueden implementarse directamente sobre hardware (GPUs y/o FPGAs), de tal forma que el procesamiento pueda incluirse en los propios sistemas físicos de visualización.

El método es una combinación del *escalado lineal simple* (presentado anteriormente) y aumento adicional del brillo en las proximidades de las zonas saturadas de la imagen (de forma similar al método presentado en el apartado anterior [MDS06]).

El algoritmo se divide en bloques funcionales, tal como se muestra en la Figura 4.5:

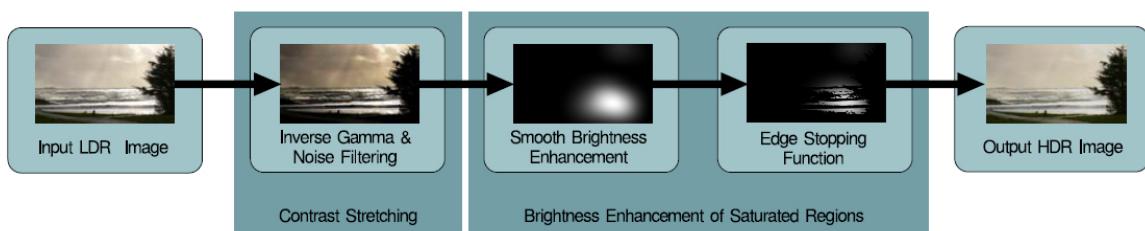


Figura 4.5: Proceso de obtención de HDR a partir de LDR utilizando el método de *Ldr2hdr: On-the-fly Reverse Tone Mapping* [RTS⁺07].

- Primero se compensa la codificación no-lineal de las imágenes LDR, aplicando el inverso de la función de codificación. Para ello se usan las curvas-gamma del monitor (habitualmente, curvas gamma de parámetro 2.2 [Com90, SAM96]). Invirtiendo esta función se obtienen luminancias de píxeles cercanas a las presentes en la escena original. Después se realiza una expansión lineal del rango dinámico, aplicando un

factor de escalado de 5000:1 (umbral bastante tradicional, pero que funciona bien para todas las imágenes). Más tarde se aplican filtros para evitar el ruido, así como la aparición de bloques y efectos indeseados (filtro Bilateral [TM98] modificado).

- Bloque 2: se realiza la función de mejora del brillo (término original: *brightness enhancement*), que se usa para aumentar el brillo en las zonas de la imagen que contengan pixels saturados (entendiendo por saturados los que estén próximos al valor máximo -255- en cualquiera de los canales RGB). Esta función es limpia en el sentido de que no produce ruidos y asegura que el aumento de brillo no soprepase contornos fuertes, además de contar con la llamada *edge stopping function* que se encarga de limitar el aumento de rango en las proximidades de bordes marcados, tal como muestra la Figura 4.6.



Figura 4.6: Aumentando el rango dinámico de las imágenes. La fotografía de la derecha es de la que quiere aumentarse el rango. La figura del centro muestra el resultado de ejecutar la función de aumento de brillo, y en la derecha se incorpora la *edge stopping function*, que previene que se aumente el rango en bordes marcados [RTS⁺07].

4.2. Diseño del test

El propósito del *reverse tone mapping* es tomar el contenido de una fotografía en bajo rango dinámico y expandir el mismo para obtener una imagen HDR, minimizando la cantidad de ruido y manchas en la fotografía resultado. ¿Alguno de los rTMO's existentes es capaz de satisfacer este requerimiento, cuando trata con imágenes incorrectamente expuestas? ¿Cuál de los algoritmos funciona mejor con cada nivel de exposición? El propósito de este Capítulo pasa por obtener respuestas a todas estas preguntas. Para ello se diseñará y llevará a cabo un experimento psicofísico que arroje algo de luz en este campo, hasta ahora inexplorado.

4.2.1. Estímulos

Se usarán los mismos que en el Capítulo anterior. Éstos son imágenes tomadas en diez escenarios distintos, capturadas por una Nikon D200 a resolución 3872 por 2592 y reescaladas a 1920 por 1080 para propósitos de visualización. Las fotografías se eligieron de forma que se cubriera un amplio abanico de iluminaciones y medios. Se realizó bracketing de exposición de cinco disparos por secuencia, desde claramente subexpuesto (etiquetadas

como 1) a altamente sobreexpuesto (etiquetadas como 5), obteniéndose un total de 50 imágenes.

En la Figura 4.7 se observa una fotografía representante de cada una de las secuencias elegidas. La Figura 4.8 muestra una secuencia completa (las cinco exposiciones distintas de una misma escena).



Figura 4.7: Estímulos: un representante de cada escena.



Figura 4.8: La secuencia completa de bracketing para la escena *sunset*, junto a sus respectivos histogramas de luminancias.

Además se dispone de una valoración estadística $(\alpha(F_i), \beta(F_i), \gamma(F_i))$ de cada fotografía F_i , con $0 \leq \alpha(F_i) \leq 1$, $0 \leq \beta(F_i) \leq 1$, $0 \leq \gamma(F_i) \leq 1$ y $\sum (\alpha(F_i), \beta(F_i), \gamma(F_i)) = 1$, siendo $\alpha(F_i)$ el porcentaje de encuestados que clasificaron como subexpuesta, $\beta(F_i)$ el porcentaje de clasificación como correcta, y $\gamma(F_i)$ el de sobreexpuesta.

Para cada una de las cinco exposiciones de cada una de las diez escenas (un total de 50 imágenes), se aplican los cuatro operadores que se detallan en la Sección *Metodología*, de forma que se obtengan cuatro imágenes para mostrar en el BrightSide DR37-P (el dispositivo de visualización HDR que se utilizará).

4.2.2. Metodología

Se realizará una comparación entre:

1. *LDR*: La imagen original LDR mostrada en el monitor HDR.
2. *Linear*: el rango de la imagen LDR original se escala linealmente para alcanzar el rango visualizable por el monitor en HDR. (Según el método descrito anteriormente en la Sección *Linear expansion*).
3. *Warwick*: rTMO basado en la inversión del reproductor de tono de Reinhard y en la construcción de un *expand map* [BLDC06] (descrito anteriormente en la Sección *Inverse Tone Mapping algorithm*).
4. *Dolby*: para transformación de imágenes "al vuelo" (presentado anteriormente en la Sección *Ldr2hdr: On-the-fly Reverse Tone Mapping*[RTS⁺07])

Nótese que el código fuente de estos algoritmos no está disponible para descarga o prueba. Por ello, hubo que contactar con Francesco Banterle y Olga Sorkine para obtener las imágenes HDR resultado de aplicar los rTMOs de *Warwick* y *Dolby* a las imágenes de entrada *LDR*. El escalado lineal (*Linear*) lo implementó el autor de este proyecto.

Estas negociaciones trataron de establecerse también con el *Ecole Polytechnique Fédérale de Lausanne (Switzerland)* y *Sharp Laboratories of America (Camas WA, USA)* para poder someter su rTMO [MDS06] a comparativa. Desafortunadamente no pudieron concluirse con éxito las negociaciones. Frente a este contratiempo, el autor de este proyecto realizó una implementación de este rTMO, con la intención inicial de que pudiera compararse directamente el algoritmo con la competencia. Sin embargo, y ante la cantidad de parámetros a estimar, de los que el autor no tenía más información que la presentada en [MDS06], se consideró injusto y de poco rigor científico realizar la comparativa en desigualdad de condiciones (en los otros casos, los autores de los rTMO's, experimentados en los valores correctos para sus algoritmos, pudieron elegir los parámetros que consideraron más oportunos a la hora de obtener resultados óptimos). Por tanto este rTMO no será sometido a examen.

En cada test los encuestados (veinte, en esta ocasión) observan simultáneamente los resultados de aplicar cada uno de los algoritmos enumerados anteriormente en una red 2x2 (conocida en el argot psicofísico como *stimulus quadruple*), tal como se muestra en la Figura 4.9. Las cuatro imágenes (en formato *.pfm*, cuyos detalles pueden observarse en el Capítulo B) se muestran en el BrightSide DR37P cedido por *MPI for Biological Cybernetics* (Tübingen), puesto que no se dispone de monitor HDR en la Universidad de Zaragoza.

Se pide a los encuestados que ordenen, usando una escala simple y familiar ('1' la mejor, '4' la peor) las cuatro imágenes de acuerdo a cuál les parece mejor, esto es, les resulta más convincente y más similar a la realidad. Se informa a los sujetos de que no existe una respuesta correcta: simplemente deben indicar su preferencia personal. Anteriores estudios ponen en evidencia que existe una fuerte correlación entre distintos criterios de juicio (como 'realismo', 'naturalidad' y 'atractivo de la imagen') [AFR⁺07, SH06]; por tanto se decidió que sería suficiente con esta pregunta.



Figura 4.9: Comparando *LDR*, *Linear*, *Warwick* y *Dolby*. Sistema de visualización: HDR BrightSide DR37-P cedido por *MPI for Biological Cybernetics* (Tübingen). A la izquierda, con las luces de la sala encendidas. A la derecha, a oscuras, tal y como se realizó el experimento.

Cada test consiste en 50 cuádruplas de estímulos mostrados en orden pseudo-aleatorio, con la restricción de que una cuádrupla y la siguiente no pueden corresponder a la misma escena. Los encuestados disponen de todo el tiempo que consideren necesario para contestar. El experimento completo consiste en tres bloques de tests. Entre cada bloque se pide a los sujetos que tomen una pequeña pausa antes de continuar.

Los 20 sujetos a quienes se realiza el test tienen visión normal o corregida a normal. La supervisión en la realización de los tests está a cargo de Roland Fleming, especialista en psicofísica y percepción, con varias publicaciones de notable importancia [FTA04, FABJ04, AFR⁺07].

Es importante resaltar algunos aspectos de interés:

1. Al haberse aplicado cada uno de los algoritmos de forma independiente (el *Linear* en Zaragoza, a cargo del autor de este proyecto; *Dolby* en Canadá, a cargo de Olga Sorkine; *Warwick* en Bristol, a cargo de Francesco Banterle) las imágenes contienen rangos lumínicos muy distintos. No tiene sentido comparar imágenes de rango 0:1 con rangos de 0:4000, especialmente cuando la intensidad global y el contraste no están entre las cosas que desean estudiarse. Así, en el experimento se han normalizado las imágenes tal como se enuncia a continuación:

- Las imágenes de *Dolby* eran las más brillantes. Los pixels con luminancia superior a 3000 se ajustaron a 3000.
- Las imágenes de *Warwick* y *Linear* se escalaron a la intensidad máxima de las de *Dolby*.
- Las imágenes *LDR* (jpeg's) se dejaron en 0:255. Sorprendentemente, estas imágenes continúan siendo favoritas en algunas ocasiones.

- Para poder realizar una comparativa de todas ellas a la vez en la pantalla, hubo que redimensionar las imágenes. Esto se hizo usando Matlab, con el comando $im = imresize(im, 0.5, 'bicubic');$
2. Dos de las imágenes enviadas por Dolby estaban corruptas (no podían leerse) en el momento de llevar a cabo el experimento, aunque fueron proporcionadas un tiempo después por Dolby. Estas dos imágenes (*computers05* y *graffitti02*), por tanto, fueron eliminadas del experimento.

4.3. Resultados

Una vez computados los resultados, se obtiene una ordenación de los rTMO's preferidos para cada fotografía. Por ejemplo, para la fotografía F_i , si los operadores preferidos son $rTMO_a, rTMO_c, rTMO_b$ y $rTMO_d$, en este orden, $\sigma(F_i)$ sería.

$$\sigma(F_i) = \{rTMO_a \ rTMO_c \ rTMO_b \ rTMO_d\} \quad (4.3)$$

Las siguientes Secciones analizan los resultados obtenidos desde varios puntos de vista.

4.3.1. Mejor rTMO global

En esta Sección se presenta una ordenación entre los cuatro métodos de transformación sometidos a examen. Se analiza cuál de ellos ha obtenido menor puntuación (y por ende mejor clasificación, puesto que una puntuación de '1' implica que el método se eligió en primer lugar. Una puntuación cuantitativamente mayor, por ejemplo '4', es cualitativamente peor, ya que indica que el método se eligió en cuarto y último lugar). Los resultados se muestran gráficamente en la Figura 4.10.

El rTMO que ha sido votado más veces como preferido es *Dolby*, seguido muy de cerca por *Linear*. Tanto *Warwick* como *LDR* han obtenido una clasificación mucho peor. La conclusión inmediata que se deriva es que, en ausencia de más datos, es mejor aplicar *Dolby* o *Linear* (que además, son mucho más rápidos en la obtención de resultados que *Warwick*, aunque lógicamente más lentos que visualizar la propia imagen *LDR*).

4.3.2. Mejor rTMO para subexposición

En esta Sección se presentará una ordenación entre los cuatro métodos de transformación sometidos a examen, cuando tratan con imágenes subexpuestas (entendiendo por subexpuesta que han sido clasificadas como tal por un 70 % o más de los encuestados en el Test 1). Los resultados se muestran gráficamente en la Figura 4.11.

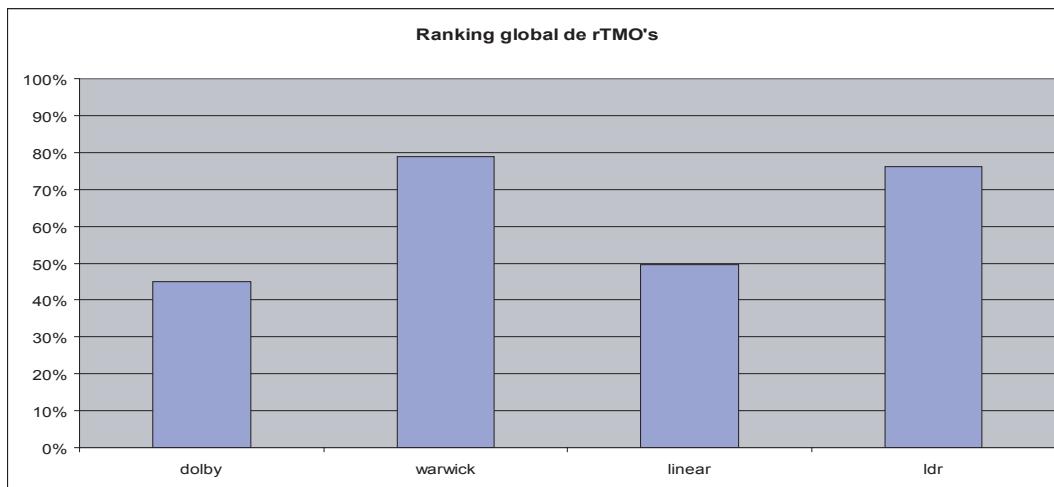


Figura 4.10: Resultados del test psicofísico: Comparativa global entre rTMO's. Se computa cuál ha tenido más votos absolutos y cuál menos. Nótese que valores inferiores en las barras indican mayor preferencia (más votos como '1', o 'mejor'). Un valor del 100 % en el eje Y quiere decir que el algoritmo ha obtenido el 100 % de valoraciones '4', esto es, que ha sido elegido siempre en último lugar.

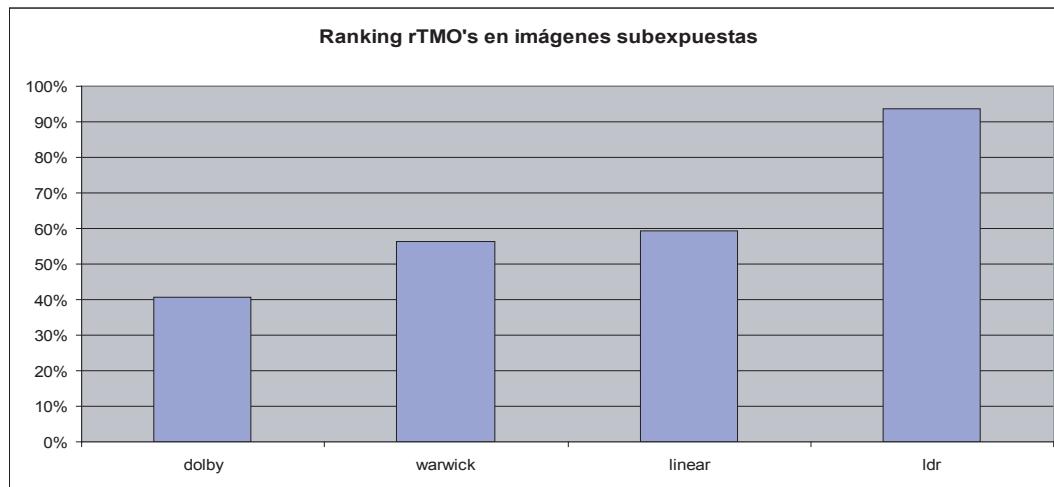


Figura 4.11: Resultados del test psicofísico: Comparativa entre rTMO's para subexposición. De las imágenes clasificadas como subexpuestas por un 70 % o más de los encuestados (en el test psicofísico 1, detallado en el Capítulo 3), se computa cuál ha tenido más votos y cuál menos. Nótese que valores inferiores en las barras indican mayor preferencia (más votos como '1', o 'mejor'). Un valor del 100 % en el eje Y quiere decir que el algoritmo ha obtenido el 100 % de valoraciones '4', esto es, que ha sido elegido siempre en último lugar.

El rTMO ganador es *Dolby*, seguido de *Linear*, *Warwick* y *LDR*.

4.3.3. Mejor rTMO para sobreexposición

En esta Sección se presentará una ordenación entre los cuatro métodos de visualización sometidos a examen, cuando tratan con imágenes sobreexpuestas (entendiendo por sobre-

expuestas las que han sido clasificadas como tales por un 70 % o más de los encuestados). Los resultados de la ordenación se muestran gráficamente en la Figura 4.12.

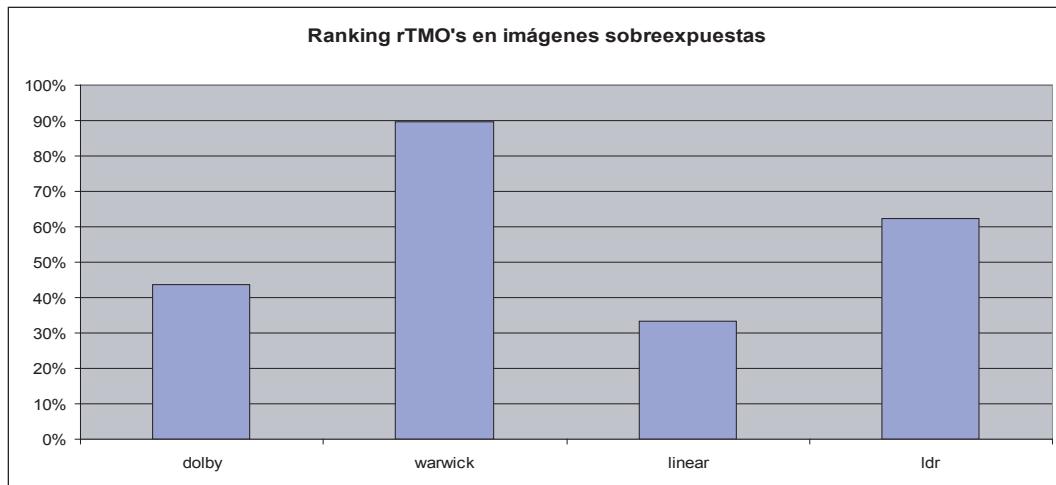


Figura 4.12: Resultados del test psicofísico: Comparativa entre rTMO's para sobreexposición. De las imágenes clasificadas como sobreexpuestas por un 70 % o más de los encuestados (en el test psicofísico 1, detallado en el Capítulo 3), se computa cuál ha tenido más votos y cuál menos. Nótese que valores inferiores en las barras indican mayor preferencia (más votos como '1', o 'mejor'). Un valor del 100 % en el eje Y quiere decir que el algoritmo ha obtenido el 100 % de valoraciones '4', esto es, que ha sido elegido siempre en último lugar.

A la vista de los resultados, el algoritmo que mejores salidas produce es *Linear*. En segunda posición se encuentra *Dolby*, seguido de *LDR* y *Warwick*.

4.3.4. Estadísticas de la imagen vs rTMO

En el Capítulo 3 se pone de manifiesto que existe una correlación importante entre datos objetivos de la imagen y la clasificación que ésta obtiene por psicofísica. En este contexto, se analizará la relación de las estadísticas de la imagen que más correlación presentan con la percepción de la exposición (media para la subexposición y porcentaje de pixels con luminancia mayor o igual a 254 en el caso de la sobreexposición.)

Para ello se observan las medias de luminancia de las imágenes que han obtenido al menos un 70 % de acuerdo en la clasificación como subexpuestas. La mayor de estas corresponde a la Fotografía *car03*, con valor 38.41 (todos estos datos se extraen del Capítulo 3.). Tomamos como umbral ese valor y obtenemos una clasificación de rTMO preferido para todas las fotografías con media de luminancias inferior a 38.41 (*car01*, *car02*, *car03*, *indoor01*, *indoor02*, *pencils01*, *computers02*, *computers03*, *computers04*, *waxes01*), tal como se observa en la Figura 4.13.

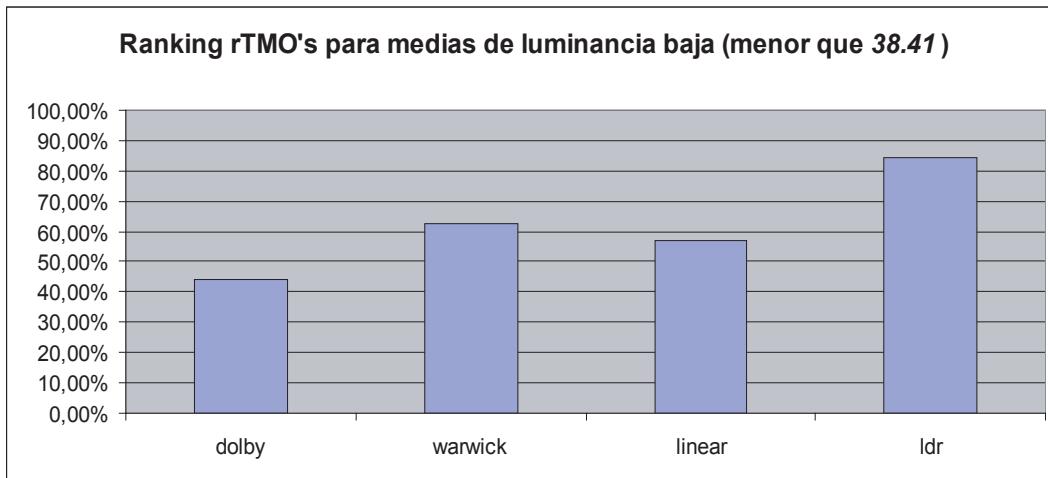


Figura 4.13: Resultados del test psicofísico: Comparativa entre rTMO's en imágenes con media lumínica baja (inferior a 38.41, umbral determinado por la Fotografía *car03*. Nótese que valores inferiores en las barras indican mayor preferencia (más votos como '1', o 'mejor'). Un valor del 100 % en el eje Y quiere decir que el algoritmo ha obtenido el 100 % de valoraciones '4', esto es, que ha sido elegido siempre en último lugar.

A continuación se estudian los porcentajes de pixels saturados de las imágenes que han sido clasificadas como sobreexpuestas por, al menos, el 70 % de los encuestados. Se toma el porcentaje mínimo (8.35 %, Fotografía *sunset03*) y se calculan los rTMO's preferidos para fotografías con un porcentaje mayor al umbral elegido (resultando el conjunto formado por *building02*, *building03*, *building04*, *building05*, *lake03*, *lake04*, *lake05*, *sunset03*, *sunset04*, *sunset05*, *graffitti04*, *graffitti05*, *strawberries04*, *strawberries05*). La Figura 4.14 muestra los estímulos percibidos como sobreexpuestos en relación al porcentaje de pixels saturados.

Los resultados de estas dos ordenaciones presentan, como era de prever, una correlación absoluta con los obtenidos en las Secciones anteriores (misma ordenación para *Fotografías subexpuestas - Medias de luminancia inferiores o iguales a 38.41* y *Fotografías sobreexpuestas - Porcentaje de pixels saturados superior o igual a 8.35 %*).

4.4. Conclusiones

La primera conclusión que se deriva a la vista de los resultados es que *Dolby* y *Linear* obtienen mejores resultados que *LDR* y *Warwick*, por norma general. Además hay que tener en cuenta que tanto *Dolby* como *Linear* son rTMO's más ligeros y por ende más rápidos que la competencia. Ambos son eficaces y eficientes, lo que les hace sumar enteros a la hora de convertirse en el algoritmo óptimo. Este resultado aporta una importante contribución, pues completa lo presentado en [AFR⁺07]. En ese estudio se comparaba el rendimiento de distintos rTMO's y se concluía que, por norma general, el mejor método era *Linear*. Sin embargo, el rTMO de *Dolby* se estaba desarrollando en el momento de la publicación del artículo anteriormente citado, por lo que no fue evaluado en ese estudio. Gracias a este test se ha podido determinar que el rendimiento de *Dolby* sobrepasa lige-

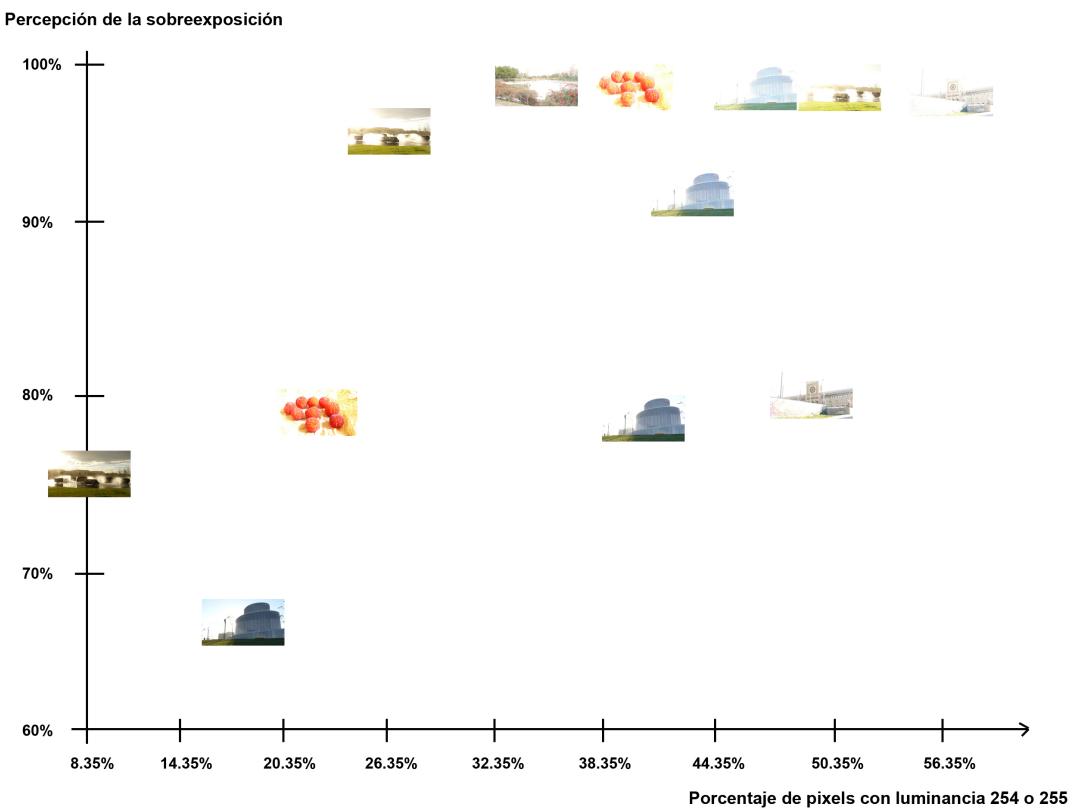


Figura 4.14: Estímulos percibidos como sobreexpuestos. El eje X muestra el porcentaje de pixels con luminancias mayores o iguales a 254. El eje Y muestra el porcentaje de encuestados que clasificaron las fotografías como sobreexpuestas.

ramente al obtenido por *Linear*.

Abordando el tema de la subexposición, el mejor rTMO es *Dolby*, seguido muy de cerca por *Linear*. *Dolby* realiza un incremento del brillo en las zonas de alta luminancia, lo que se traduce en imágenes más claras, suavizando el efecto de oscuridad y pérdida de detalle ocasionado por la subexposición. En tercera posición se encuentra *Warwick*, que se beneficia del aumento de brillo en las zonas luminosas (generado gracias al *expand map*), para superar en rendimiento al método *LDR*, que mantiene la oscuridad característica de las fotografías subexpuestas. Además, el análisis de la preferencia de rTMO's asociado a estadísticas de primer orden de la imagen (más en concreto, de la media de luminancia de píxeles, que presenta una alta correlación con la percepción de subexposición) corrobora estos resultados.

Para las fotografías sobreexpuestas (y que, tal como se demostró en el Capítulo 3, tienen un porcentaje alto de píxeles con luminancias mayores o iguales a 254), se prefiere *Linear* a *Dolby*, pues este último método tiende a agudizar el problema de exceso de brillo en las zonas sobreexpuestas. Este es también el motivo de que se prefieran las Fotografías *LDR* a las obtenidas al aplicar *Warwick*.

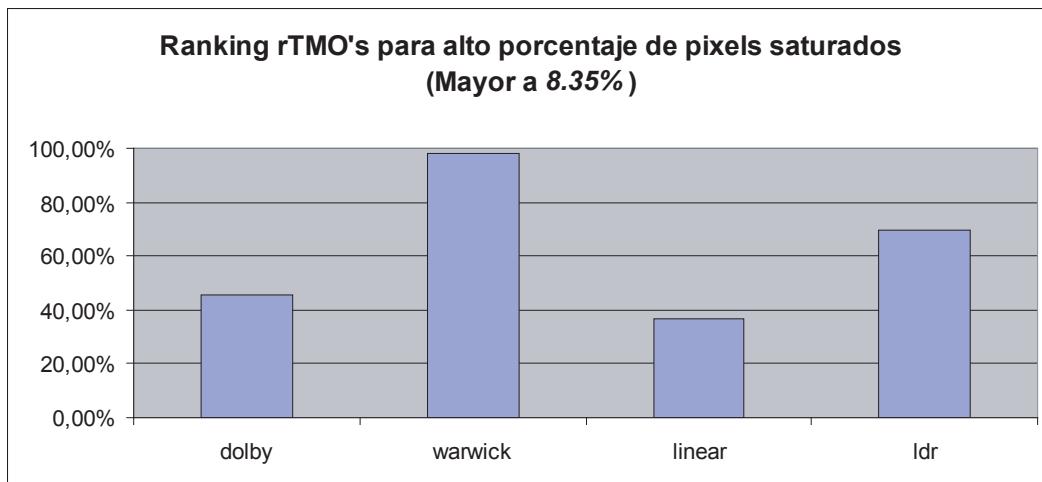


Figura 4.15: Resultados del test psicofísico: Comparativa entre rTMO's en imágenes con porcentaje de pixels saturados alto (superior a , umbral determinado por la Fotografía . Nótese que valores inferiores en las barras indican mayor preferencia (más votos como '1', o 'mejor'). Un valor del 100 % en el eje Y quiere decir que el algoritmo ha obtenido el 100 % de valoraciones '4', esto es, que ha sido elegido siempre en último lugar.

Era de esperar que las imágenes *LDR* fueran consideradas peores que el resto, pues no se benefician del aumento de rango dinámico proporcionado por la tecnología HDR. La baja valoración obtenida para el rTMO *Warwick* se debe a la presencia de ruidos de alta frecuencia, así como de objetos indeseados en las imágenes resultado. Incluso aunque las imágenes obtenidas por el método de *Warwick* sean, en ocasiones, las que mejor apariencia tienen, la presencia de estos ruidos puntuales ocasiona el rechazo obtenido. Además, el incremento del brillo en las zonas muy saturadas perjudica los resultados obtenidos para fotografías sobreexpuestas. Hay que decir, en favor de este último método, que los resultados puede verse influenciados por la normalización aplicada a las imágenes, ya que los histogramas de *Warwick* tienden a estar más desplazados a la derecha que el resto.

Hay que notar que, pese a que el método *Linear* no sea a menudo el mejor, es el que menos probabilidad tiene de introducir ruido y objetos indeseados en la imagen resultado. Esto le hace situarse por encima de la competencia en muchas ocasiones, pues los sujetos encuestados son muy sensibles a la aparición de este tipo de ruidos.

Capítulo 5

Diseñando un nuevo rTMO

El Capítulo 3 presenta una clasificación de los estímulos en (1)subexpuestos, (2)correctamente expuestos y (3)sobreexpuestos, así como el establecimiento de hipótesis de causalidad acerca de la percepción de la exposición (¿Qué hace que una imagen se en-globe en uno de los grupos anteriores?).

El Capítulo 4 presenta una comparativa entre los rTMO's existentes cuando tratan con imágenes que pertenecen a cualquiera de los tres grupos anteriores. En él se detalla cuál de los rTMO's comparados funciona mejor con cada uno de los tipos de imagen (en cuanto a exposición se refiere) y se enuncian las causas que motivan estos resultados.

En este Capítulo se utiliza el conocimiento obtenido en anteriores Capítulos para plantear el diseño de un nuevo rTMO que se beneficie de las bondades de los existentes, evitando (en la medida de lo posible) los defectos observados en cada uno de ellos.

5.1. Análisis de la fotografía LDR a transformar

El primer paso del nuevo rTMO pasa por analizar minuciosamente las imágenes a transformar. Más en concreto, se tratará de determinar, mediante estadísticas de primer orden de las imágenes, si la fotografía a expandir está en el grupo de las subexpuestas, sobreexpuestas o correctamente expuestas. Tal y como se enunciaba en el Capítulo 3, la percepción de la exposición fotográfica es subjetiva y variable de un individuo a otro. Sin embargo, existen indicadores objetivos que muestran correlaciones importantes con la percepción de la exposición.

Más en detalle, las fotografías subexpuestas tienen medias de luminancia de píxeles bajas. El umbral de subexposición, calculado en el Capítulo anterior, se fija en media de luminancia igual a 38.41. Las fotografías percibidas como sobreexpuestas cuentan con un número elevado de píxeles con luminancias mayores o iguales a 254 (se utiliza este valor como umbral de saturación, tal y como se hace en [RTS⁺⁰⁷]). Una fotografía sobreexpuesta tiene un porcentaje de pixels saturados superior o igual a 8.35 %, tal como se calcula en

el Capítulo 4.

En ese mismo Capítulo se pone de manifiesto que la clasificación de las imágenes utilizando exclusivamente datos objetivos de la imagen no es *perfecta*. Se haría necesario analizar semánticas de alto nivel, como por ejemplo el reconocimiento de elementos de la imagen, para poder establecer una clasificación más precisa de los estímulos. El cerebro humano es capaz de distinguir patrones; esto es, es capaz de discernir si una mancha gris en la imagen corresponde a un oso polar o un objeto sobreexpuesto. Una vez reconocido el patrón y analizadas las fuentes de iluminación, el cerebro determina rápidamente si la exposición de la imagen le resulta adecuada o no. Sin embargo, pretender que un ordenador sea capaz de llevar a cabo esta tarea resulta utópico, a día de hoy. Por tanto, a la hora de analizar la fotografía, deben aplicarse criterios que una máquina sí sea capaz de determinar por sí misma (sin intervención humana), como son las estadísticas de primer orden de una imagen.

Para cada imagen entrada se analizarán los dos aspectos que más correlación presentan con la percepción de la exposición: la media de valores lumínicos de los píxels que la componen y el porcentaje de los mismos con luminancia mayor o igual a 254. Nótese que la obtención de estas estadísticas puede implementarse en FPGA's y, por tanto, sobre el propio hardware de visualización o reproducción (televisores, monitores, reproductores de DVD y otros).

5.2. Aplicación del rTMO óptimo

Una vez observadas estas estadísticas, se aprovecha la información obtenida en el Capítulo 4. De él se extraen varias reglas, que serán aplicadas a las fotografías a expandir por el nuevo rTMO:

- Cuando la fotografía tiene medias de luminancia por debajo del umbral establecido (38.41), el algoritmo que mejor funciona y por tanto se aplicará es *Dolby*.
- Cuando la fotografía tiene un 8.35 % o más de pixels con luminancia mayor o igual que 254, el algoritmo que obtiene mejores resultados es *Linear*.
- Raramente se dan fotografías que cumplan los anteriores criterios a la vez (corresponderían a imágenes con contrastes altísimos). Sí son frecuentes las imágenes que no cumplen ninguno de estos criterios (de hecho, la mayoría de las que no los cumplen se perciben como correctamente expuestas). A las fotografías que no cumplen ninguno de los criterios anteriores, o cumplen los dos, se les aplicará el rTMO *Dolby*, por ser el que consigue mejores resultados en general.

Además, los algoritmos *Linear* y *Dolby* únicamente necesitan algunos parámetros dependientes del dispositivo de visualización en que se mostrarán los resultados (resolución o amplitud del rango dinámico deseado). Por tanto, la aplicación de ambos rTMO's puede

realizarse "al vuelo", siendo apto para la transformación de secuencias de video. Es importante notar, además, que estos métodos pueden implementarse directamente sobre hardware ([RTS⁺07]) y, por ende, estar situados en el propio dispositivo de visualización.

5.3. Normalización de resultados

Tal y como se presentaba en el Capítulo 4, los resultados obtenidos por *Dolby* y *Linear* deben estar normalizados a una misma luminancia máxima. De ese modo se consigue obtener secuencias de imágenes coherentes, sin que se aprecien cambios de brillo notables derivados del uso de uno u otro algoritmos. En el Capítulo 4 se utilizó el valor de luminancia máxima 3000, obteniendo buenos resultados, por lo que en este caso se usa este mismo umbral.

5.4. Conclusiones

Se ha diseñado un rTMO que obtiene buenos resultados para todo tipo de imágenes, independientemente de si presentan o no correcta exposición. Este algoritmo es capaz de trabajar en tiempo real, y no presenta diferencias notables entre un fotograma y el siguiente, salvo que las diferencias estén en los fotogramas originales. Además es posible implementarlo sobre hardware directamente. La Figura 5.1 muestra, de modo esquemático, las fases en que se divide el mismo.

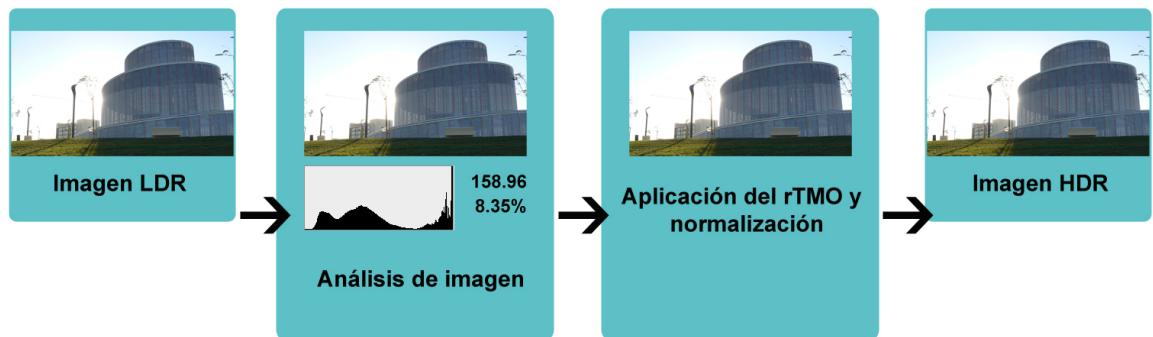


Figura 5.1: Pasos del nuevo rTMO: análisis, aplicación de rTMO en función al análisis y normalización de salida.

Capítulo 6

Problemas, conclusiones y trabajos futuros

6.1. Conclusiones y trabajo futuro

El *reverse tone mapping* (o mapeado de tono inverso) es un proceso para el cual no existe una solución definitiva. Dada la creciente disponibilidad de dispositivos de visualización HDR, la cuestión de cómo mostrar la inmensa cantidad de contenido LDR legado se ha convertido en un tema candente. Según Google Scholar, existen más de 900 artículos de investigación relacionados de forma directa con el proceso del mapeado de tono, que concluyen en varias docenas de algoritmos distintos [MS08]. Cada uno de estos algoritmos resuelve el problema de forma parcial o incompleta. Con este proyecto se trata de evitar que se produzca una situación similar para el problema del mapeado de tono inverso, puesto que se pretende desarrollar un único algoritmo que funcione bien con todas las imágenes. Existen rTMO's capaces de obtener buenos resultados [AFR⁺07], pero todos ellos asumen que las imágenes de entrada son perfectas. Dado que el volumen de fotografías perfectas representa una ínfima parte del conjunto de fotografías disponibles, debe encontrarse una solución que sea capaz de trabajar de forma aceptable con imágenes de cualquier tipo. Una fotografía puede ser imperfecta por muchos motivos: presencia de ruido, compresión, desenfoque, incorrecta exposición... Este proyecto se centra en solucionar éste último problema: el relativo a la incorrecta exposición de la instantánea.

El diseño del nuevo algoritmo (presentado en el Capítulo 5) ha sido construido sobre la base de conocimiento aportada por los tests psicofísicos de los Capítulos 3 y 4. En estos tests se utilizan cincuenta fotografías y un grupo de encuestados para obtener información acerca de la percepción de la exposición y de las preferencias de uso de los distintos rTMO's existentes. Es obvio que de haber contando con mayor número y variedad de fotografías, así como con un mayor número de encuestados aportaría un nivel de conocimiento mayor al experimento. Sin embargo, tanto la realización de los tests como la extracción y análisis de resultados habrían sido muy costosos, excediendo los propósitos de este proyecto.

En el Capítulo 3, además de obtener una taxonomía de los estímulos, se establecieron algunas hipótesis relacionadas con la percepción de la exposición. Gracias a ello se abre un campo de investigación interesante, en el que deben verificarse las hipótesis planteadas, así como de-correlar los indicadores estadísticos utilizados o emplear otros más precisos (¿conviene analizar el histograma por canales, y no únicamente de forma global?).

Un análisis por individuos resultaría interesante desde el punto de vista de la compresión de la percepción de la exposición. Deberían analizarse las diferencias de comportamiento entre individuos, así como las diferencias entre las valoraciones de un individuo consigo mismo. ¿La valoración de la exposición cambia de un individuo a otro? Podría obtenerse una media de valoraciones de cada encuestado:

- Se parte del valor 0.
- Si el individuo_i valora la fotografía F_j como subexpuesta, se suma -1 , si valora como correcta 0 y si lo hace como sobreexpuesta, $+1$.
- Sumando la valoración de todas las fotografías se obtiene la media de valoraciones por individuo. Valoraciones cercanas a 0 indican que éste tiende a clasificar las fotografías como correctas, mientras que cuanto más cercano sea ese valor a ± 1 , mayor será la tendencia a clasificar como sub o sobreexpuestas.

También es interesante analizar la fiabilidad de las respuestas, esto es, valorar en qué medida varían las contestaciones entre ejecuciones del test para un mismo individuo. ¿Se observan tendencias generales? ¿Hay alguna secuencia en que los participantes contesten cosas muy distintas entre ejecuciones del test? ¿Existen usuarios *más fiables* (esto es, que contesten de forma más consistente) que otros? ¿Existe una correlación entre la *fiabilidad* del usuario y el nivel de conocimiento fotográfico que tiene?

Además, en el Capítulo 4 se estudia la preferencia de los distintos rTMO's en función de la fotografía. Sin embargo, ¿existe algún otro elemento común a todas las fotografías con el mismo rTMO preferido? Esto es, si se realiza una ordenación de las fotografías por rTMO preferido y se analizan (mediante un motor estadístico) los resultados obtenidos, ¿existen patrones comunes? ¿Tal vez el color predominante en una fotografía condiciona la elección de un rTMO u otro? ¿Quizás influye el tipo de iluminación? ¿O son los elementos presentes en la escena? Como advierte el lector este es un campo con grandes posibilidades.

El rTMO presentado se vale del conocimiento obtenido en los test anteriores. Así, para una imagen que ha sido clasificada en el test 1 con la tupla $(\alpha(F_i), \beta(F_i), \gamma(F_i))$, se busca el $\max \alpha(F_i), \beta(F_i), \gamma(F_i)$ y se aplica el algoritmo de *Dolby* o *Linear* en función de este valor. Sin embargo, puede que el rTMO óptimo haya de determinarse utilizando los tres valores de la tupla, y no únicamente el máximo. Es necesario un estudio más completo para poder esclarecer esta cuestión.

Además de las mejoras presentadas, la principal línea de ampliación del trabajo consiste en estudiar el resto de factores que pueden convertir a una fotografía en imperfecta

(compresión, la presencia de ruido, el desenfoque) y modificar el rTMO presentado de forma que sea capaz de resolver estos problemas.

6.2. Problemas encontrados

Un problema que se planteó al inicio del proyecto fue la imposibilidad de disponer de un monitor HDR en la Universidad de Zaragoza. Esto ha fomentado la colaboración con distintas universidades europeas. Más en concreto, el Max Plankt Institut ha demostrado entusiasmo en participar en este proyecto, considerándolo de gran interés para la comunidad científica. Para ello se ha contado con el apoyo de diversos profesionales de reconocido prestigio en el campo de la investigación de Gráficos por Ordenador (como Karol Myszkowski [Kar]), así como en el campo de la percepción y la psicología (Roland Fleming [Rol]). Este mismo centro universitario invitó a una semana de estancia en sus instalaciones al autor de este proyecto, de forma que pudiera disponer de un profundo conocimiento de la tecnología HDR y la problemática existente con el mapeado de tono inverso. Asimismo se han tenido que establecer conversaciones con la *University of British Columbia, Dolby Canada, University of Bristol, Ecole Polytechnique Fédérale de Lausanne (Switzerland)* y Sharp Laboratories of America para la obtención de resultados de aplicar sus respectivos rTMO's a las imágenes elegidas como estímulos. Todas estas conversaciones han supuesto algunos retrasos en la planificación del proyecto, como era de esperar.

Otros problemas han surgido de el BrightSide DR37P en sí mismo, como por ejemplo la inestabilidad del dispositivo (una misma zona de la pantalla presenta, para un mismo fotograma, valores de brillo diferentes en función de cuánto tiempo lleve el monitor encendido). Las fluctuaciones son más acusadas cuando el dispositivo se sobrecalienta, lo que sucede con relativa facilidad. Además, las imágenes tardan bastante tiempo en cargarse, lo que supone un reto a la paciencia del encuestado.

Apéndice A

Conceptos fotográficos

A.1. Introducción

En fotografía digital, un campo de desarrollo abierto es el aumento del alto rango dinámico. En este Anexo se tratarán algunos de los aspectos de esta nueva corriente fotográfica, al tiempo que se detallan algunos conceptos fotográficos relevantes para este Proyecto.

La nueva tendencia en el mundo de la fotografía viene a suplir algunos de los problemas de la fotografía tradicional, como son:

- Las escenas reales contienen rangos lumínicos que varían en una proporción enorme (desde la luz del sol a la oscuridad más absoluta).
- Durante cientos de años, el medio donde se mostraban las fotografías (inicialmente papel fotográfico y más adelante, monitores convencionales) ha estado limitado a rangos lumínicos insuficientes para mostrar correctamente la luminosidad de escenas reales. ([HDRa]).
- Nos encontramos, por tanto, ante un problema evidente: ¿cómo representar valores de luz de una escena real usando un conjunto mucho más limitado de valores, que es el que admite el medio?

Un ejemplo de este fenómeno es la Figura A.1 , que se muestra a continuación. En la escena real encontramos variaciones enormes de intensidad lumínica entre los puntos del exterior de la ventana (muy iluminados) y el sofá (sin apenas iluminación). Esta información ha de limitarse, al tomar la fotografía, al rango de valores que admite el medio. Por tanto, resulta imposible mostrar de forma correcta, y en una misma fotografía, las zonas muy iluminadas (la ventana) y las oscuras (el sofá) sin que ninguna de ellas pierda detalle.



Figura A.1: La zona de la ventana está sobreexpuesta y la de la del sofá, subexpuesta.

A.2. Exposición

En el caso anterior se dice que una parte de la fotografía (lo que se ve por la ventana) está sobreexpuesta (ya que el medio fotográfico se ha expuesto a la luz durante un tiempo excesivo) y otra (el sofá) subexpuesta (exposición insuficiente a la luz del medio fotográfico).

La exposición es una combinación del lapso de tiempo y el nivel de iluminación que recibe el material fotosensible. En la cámara, el tiempo de exposición se controla con la velocidad de obturación y el nivel de iluminación, con la apertura (de la lente). Velocidades de obturación bajas (exponer el medio un periodo largo de tiempo) y mayores aperturas (admitiendo más luz) producen exposiciones altas.

La exposición para una fotografía viene determinada en última instancia por la sensibilidad (velocidad) del medio utilizado y se mide utilizando una escala publicada por la International Organization for Standardization (ISO). Los medios más rápidos requieren menos exposición a la luz y tienen un valor ISO más alto.

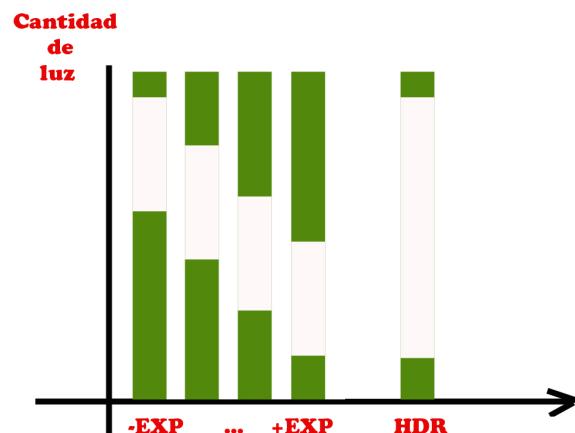
La subexposición puede deberse a la colocación incorrecta del mando de sensibilidades, a un cálculo incorrecto en situaciones difíciles de iluminación, a suciedad acumulada en el objetivo o a la no compensación de un filtro denso. Los negativos subexpuestos son muy transparentes y sin detalle, las copias planas y carentes de densidad y las diapositivas muy oscuras.

Exposiciones bajas dan como resultado fotografías en las que los puntos muy brillantes se ven nítidos, sin pérdida de detalle (y las zonas oscuras en sombra, sin detalles). Exposiciones altas dan como resultado fotografías en que las zonas muy oscuras se observan



Figura A.2: Misma escena, distintas exposiciones (de más baja a más alta, de arriba abajo).

correctamente iluminadas (y las zonas muy iluminadas sobre-saturan el medio) (ver Figura A.2). Sin embargo, en muchas ocasiones no podemos encontrar un valor de exposición que consiga mantener todas las zonas de la fotografía correctamente iluminadas. La tecnología HDR es capaz de suplir esta carencia de la fotografía tradicional, como se observa en la Figura A.3.



Cada fotografía HDR engloba el rango lumínico de varias exposiciones LDR (de fotografía tradicional).

Figura A.3: Los valores verdes representan las zonas que no tendrán una correcta exposición (se verán muy claras o muy oscuras).

Las cámaras de hoy en día son capaces de aproximar una exposición correcta de forma automática, para la gran mayoría de las escenas a fotografiar. Se conecta el exposímetro al obturador, al diafragma o a ambos mecanismos. Tiene cuatro modalidades principales [rsf] :

1. **Prioridad a la apertura:** el fotógrafo selecciona la apertura ¹, y la cámara, de forma automática, la velocidad de obturación ². Este modo se utiliza cuando es importante controlar la profundidad de campo ³.



Figura A.4: Imágenes con mayor profundidad (izda.) y menor profundidad de campo (dcha.)



Figura A.5: Variación de la profundidad de campo con la apertura del diafragma

2. **Prioridad a la velocidad:** el fotógrafo ajusta la velocidad, y la cámara, en función de esa velocidad y de las condiciones de luz, calcula la apertura del diafragma. Este modo encuentra su sentido en la fotografía deportiva y de acción (Figura A.6).

¹ La apertura de un sistema óptico es aquella que determina el ángulo del cono de un haz de rayos que se enfoca en el plano de imagen, [Wika]

² Velocidad a la que se abre y cierra el obturador de un dispositivo fotográfico. El inverso al periodo de tiempo durante el cual está abierto éste. [Wike]

³ La profundidad de campo es el rango de distancias reproducidas en una fotografía donde la imagen es aceptablemente nítida comparada con el plano más nítido de la misma. [Wikb]



Figura A.6: En fotografía deportiva se utiliza la prioridad a la velocidad, para capturar una imagen nítida.

3. **Programada (automática)**: la cámara determina la velocidad y apertura. Es el modo más útil cuando no se dispone de tiempo para preparar la toma de la instantánea. Todas las cámaras digitales de hoy en día admiten este modo. Utilizan la información proporcionada por el exposímetro (Instrumento que mide la intensidad luminosa y determina los valores de abertura y velocidad de obturación para obtener una exposición correcta). Los exposímetros incorporados a la cámara suelen dar preferencia a alguna parte determinada del motivo, por lo general al centro del mismo.



Figura A.7: Exposición automática (usando Sony CyberShot DSC-P100)

4. **Manual**: el fotógrafo decide ambos ajustes (velocidad y apertura).

Sin embargo, el concepto de exposición correcta es muy subjetivo, como se ve más adelante, puesto que una escena puede tener múltiples exposiciones *correctas*, dependiendo del efecto deseado por el fotógrafo.

A.3. Histograma

Otro concepto interesante, de cara a los test realizados, es el de histograma. Histograma es una representación gráfica de una variable en forma de barras, donde la superficie de cada barra es proporcional a la frecuencia de los valores representados. En el eje vertical se representan las frecuencias de aparición, y en el eje horizontal los valores de luminancias (del negro al blanco).

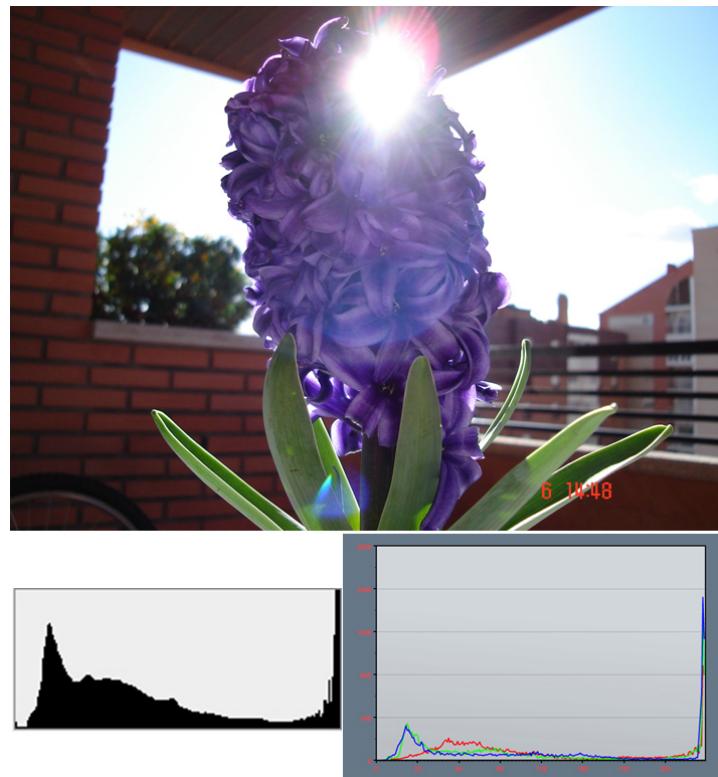


Figura A.8: Histogramas de la fotografía: a la derecha el global, de Photoshop, y a la izquierda por canales, implementado en este estudio.

El histograma global de una imagen se calcula, para cada píxel, como se indica en la Fórmula A.1, siendo R el valor de rojo, G el de verde y B el de azul [ITU]:

$$L(Pixel_{ij}) = 0.213R + 0.715G + 0.072B \quad (A.1)$$

Además, puede realizarse un histograma por canales. En este se computa cada canal de forma independiente, obteniendo tres histogramas diferentes (*Rojo*, *Green*, *Blue*). También suele representarse como un único histograma con tres funciones diferentes.

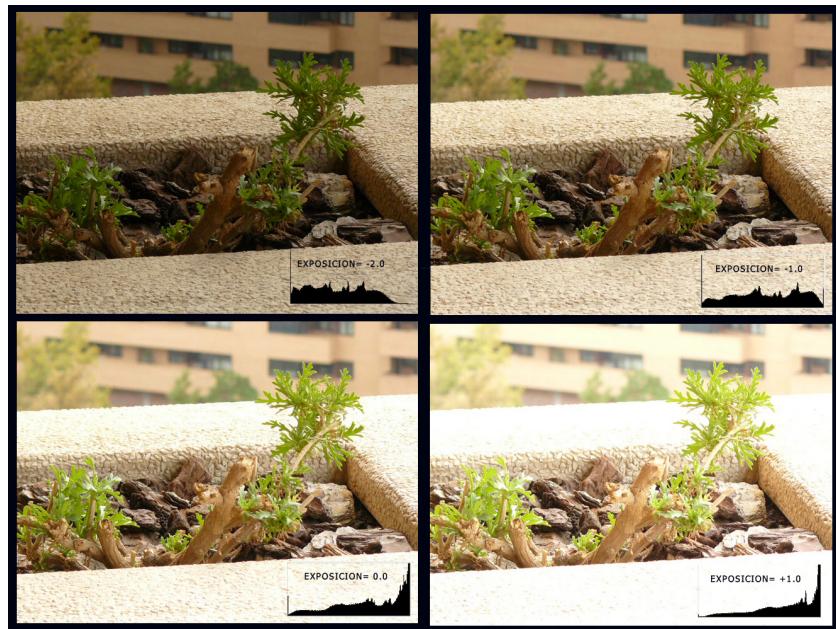


Figura A.9: Tipificación usando el histograma: Si los pixels están desplazados en la gráfica hacia la izquierda, predominan los tonos oscuros. O bien la foto está subexpuesta o bien estamos ante una imagen clave baja. Si la mayoría de los pixels están a la derecha predominan los tonos claros. O la foto está sobreexpuesta o es una imagen de clave alta. Si los pixels se amontonan en la parte central y desaparecen de los extremos significa que la imagen tiene poco contraste. Si, por el contrario, se amontonan en ambos lados, dejando la zona casi desierta, es muy probable que haya un exceso de contraste.

A.4. Soluciones clásicas al problema de la limitación del rango dinámico.

Los pintores clásicos fueron conscientes del problema de la limitación del rango de luminancias que el medio admitía. Idearon pequeños trucos que hacen que nuestro cerebro tenga la ilusión de observar una escena realista, con amplio rango dinámico.

Uno de estos famosos trucos fue el que usó El Greco en *La Agonía en el Jardín*, de 1597 [ElG]. Tal como se muestra en la Figura A.10, el Greco usó colores saturados al lado de tonos opuestos para dar sensación de más rango dinámico (o contraste apreciado [Mey]). Además, pintaba de blanco o negro los contornos de las figuras para incrementar el contraste percibido.

A.5. El rango dinámico en la fotografía clásica

En una fotografía intervienen el dispositivo de captura (cámara), el medio de almacenamiento (p. ejemplo, película) y un dispositivo de visualización (papel o pantalla). El rango dinámico de cada etapa (captura, almacenamiento y salida) juega un papel fundamental en la calidad de los resultados. En general, las tecnologías con mayor rango



Figura A.10: Los pintores clásicos ya conocían el problema de la limitación del rango dinámico. (La agonía en el Jardín, de El Greco).

dinámico producen mejores resultados. Sin embargo, cuando el rango dinámico de la escena es demasiado grande para alguna de las etapas del proceso, se debe sacrificar algo (perder detalle en las zonas iluminadas, o en las zonas de sombras). Los fotógrafos deben conocer y trabajar con las limitaciones de cámaras, sistemas de almacenamiento y dispositivos de salida.

A.6. El rango dinámico en la fotografía en color

Los negativos en color tienen menor rango dinámico que las películas en blanco y negro. Las primeras películas en color tenían muy poco contraste, así que los fabricantes añadieron más capas (cada color era separado en dos: una capa de alta sensibilidad y una capa de baja, tal como se detalla en la Figura A.11).

Un método para extender el rango dinámico en la fotografía en color es utilizar fotografía en escala de grises junto con filtros de color. Es necesario tomar tres fotografías diferentes -una con filtro rojo, otra con verde y otra con azul- y componer estas tres imágenes. Un ejemplo de este resultado lo encontramos en la obra de Prokudin Gorski [Gor] que data de 1915 (Figura A.12).

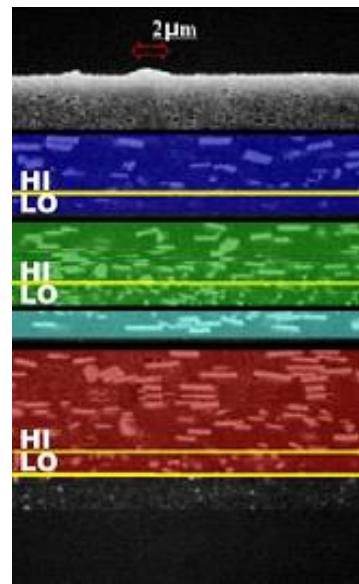


Figura A.11: Capas del negativo a color.



Figura A.12: Obra de Prokudin Gorskii en 1915, usando filtros de color.

Apéndice B

HDR

B.1. Introducción

Habitualmente se denomina rango dinámico al número de valores diferentes que puede adquirir, medir o representar un determinado medio o dispositivo. En este sentido, el rango dinámico que es capaz de representar una imagen digital convencional ronda los dos órdenes de magnitud utilizando un byte por píxel para cada uno de los tres canales en que se muestra habitualmente la luz (RGB, o *red green blue*). Esta forma de almacenar las imágenes suele ser suficiente teniendo en cuenta que la mayoría de los dispositivos disponibles para visualizar dichas imágenes no suele superar los dos órdenes de magnitud en su capacidad para variar la intensidad lumínica.

Sin embargo, el sistema de visión humano puede distinguir un rango dinámico de, aproximadamente, cinco órdenes de magnitud de forma simultánea, y de casi diez si se tiene en cuenta que es capaz de adaptarse a las condiciones de iluminación. En el mundo real, el rango lumínico de valores es mucho mayor. El sol de media tarde es aproximadamente cien millones de veces más brillante que la luz nocturna de las estrellas [Wan95]. Se puede observar una representación gráfica de estas magnitudes en la Figura B.9.



Figura B.1: Las fases del proceso fotográfico: captura, almacenamiento, edición y salida.

Durante la próxima década, la industria de la imagen realizará la transición a la tecnología de alto rango dinámico (HDR). Este cambio afectará a dispositivos de captura,

almacenamiento, edición y salida. (Figura B.1). En este Apéndice observaremos los cambios que se producirán en cada una de las etapas del proceso.

B.2. Captura

Las cámaras de hoy en día tienen una amplia resolución, por lo que el próximo campo de mejora para el sector será la calidad de los píxeles, más que el número de estos. Este cambio es ya una realidad.

Por ejemplo, la S3-Pro (fabricada por Fuji) tiene un chip SuperCCD con sensores de alta y baja sensibilidad para cada píxel [DP-] para aumentar el rango dinámico (de forma análoga al método usado en la película fotográfica. Ver Apéndice A), tal como se observa en la Figura B.2. Así, aunque el chip produce fotografías de menor resolución, éstas tienen mayor rango dinámico.

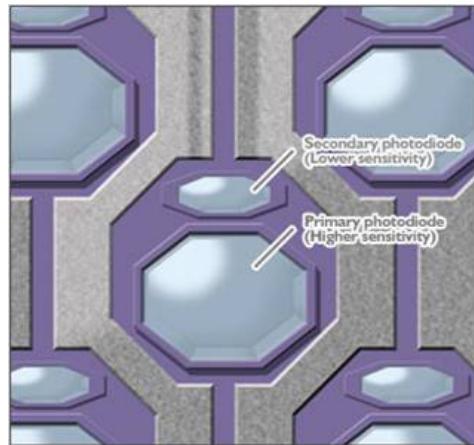


Figura B.2: Sensores en chips fotográficos (cuentan con fotodiodos de alta y baja sensibilidad).

Otra alternativa para aumentar el rango dinámico es juntar la información de varias fotografías. La técnica, que fue mostrada en el artículo [DM97] de Paul Debevec [Deb] en *SIGGRAPH 97*, consiste en tomar múltiples exposiciones de una escena y combinarlas para crear una única imagen de alto rango dinámico. Esta técnica la incorporan, a día de hoy, productos como Adobe Photoshop CS 3[Pho]. Sin embargo, este método es difícil de aplicar a imágenes borrosas o en movimiento (el mínimo movimiento puede provocar que posteriormente no "encajen" las distintas tomas que se hayan realizado [JLR07]).

Capturando la misma escena con diferentes tiempos de exposición cada imagen de la secuencia tendrá diferentes píxeles debidamente expuestos, y otros subexpuestos o sobreexpuestos. Sin embargo, cada píxel estará correctamente expuesto en uno o más disparos de la secuencia. Bajo el supuesto de que el dispositivo de captura es perfectamente lineal, cada exposición puede llevarse al mismo dominio, dividiendo la intensidad lumínica

de cada píxel por el tiempo de exposición de la imagen. Una vez *normalizadas* todas las instantáneas, la imagen resultante se obtiene realizando la media de intensidades de cada uno de los píxeles correctamente expuestos (entendiendo por correctamente expuestos que no superen ciertos umbrales de intensidad lumínica). El resultado es una imagen de alto rango dinámico.

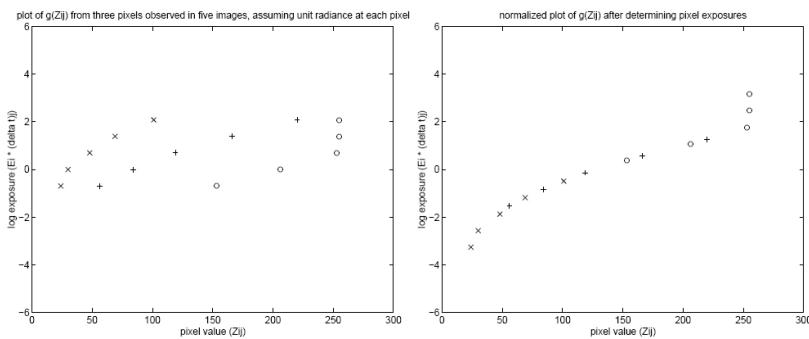


Figura B.3: Muestra de algunos píxeles antes (izquierda) y después (derecha) de la normalización de Paul Debevec.

Desde el punto de vista de la mayoría de los clientes, HDR es una tecnología que aportará a sus fotografías domésticas un mayor nivel de detalle en sombras y luces. Otra ventaja clara de este avance será que se eliminará la necesidad de usar flash. Cualquiera con una mínima experiencia en fotografía sabe que el flash, utilizado como iluminación principal, elimina la sensación de volumen en la captura, empobreciendo la instantánea. Los fotógrafos profesionales también se beneficiarán de HDR. Con la nueva tecnología en cámaras se abre un mundo de posibilidades.

B.3. Almacenamiento

Todos los formatos de almacenamiento de imágenes tienen limitaciones. Algunos, como JPEG [JPEa] y GIF [GIF] proporcionan 8 bits por cada canal de color (24bits por píxel en total). La mayoría de los formatos de 8 bits, como JPEG o GIF, usan *perceptual mapping*. Podemos encontrar una explicación más detallada del mapeado perceptual en [Kor] y [Gam]. Esto hace que los formatos de 8 bits sean ideales para almacenar fotografías con rango dinámico limitado (fotografías tradicionales). Sin embargo, estos formatos no son capaces de almacenar fotografías en alto rango dinámico.

Existen nuevos formatos (JPEG2000 [JPEb], RAW [RAW] y PNG [PNG]) que almacenan 16 bits por cada canal de color. Sin embargo, no dan soporte a la gran variedad de colores que existen, pues asignan el valor 0 al negro y 65535 al blanco. Estos valores son claramente inferiores a la variedad de luminancias existentes en la naturaleza.

Uno de los formatos más utilizados hoy en día para almacenamiento HDR es OpenEXR [Ope], desarrollado por Industrial Light & Magic. También existe TIFF-SGI, Floating Point TIFF, RGBE (de Radiance) y *Portable Float Map* o PFM [PFM]. Estos formatos almacenan 32 bits por canal de color, con lo que se consigue almacenar con precisión la luminancia real de las escenas, por encima incluso de la que son capaces de mostrar los monitores de uso doméstico actuales.

B.3.1. PFM (*Portable Float Map*)

Hay muchos formatos conocidos como PFM, y ninguno de ellos está oficialmente documentado. [PFM]. El formato que se describe aquí es el que se utilizará en este proyecto.

Una imagen en PFM es un conjunto de bytes. Este conjunto está dividido en cabecera y cuerpo (*raster*). La cabecera está compuesta por tres "líneas" consecutivas de texto ASCII, separadas por el carácter espacio.

La primera línea (o línea de identificación) contiene los caracteres "PF" o "Pf", que describen si la imagen es en color (PF) o en escala de grises (Pf).

La segunda línea (o línea de dimensión) contiene dos enteros positivos (decimal), separados por un espacio en blanco. El primer número es el ancho de la imagen y el segundo, la altura. La unidad de ambos es el píxel.

La tercera línea (o línea de escalado-almacenamiento) indica la escala y el modo de almacenar los píxeles (*little endian* o *big endian* [Wikd]). Esta línea consiste en un valor decimal no nulo, no necesariamente un entero. El valor absoluto del número indica la escala. Si es menor que cero el PFM se almacena como *little endian*. En caso contrario, como *big endian*.

El cuerpo del PFM es una secuencia de píxeles, uno tras otro, sin delimitadores de ningún tipo entre ellos. Están ordenados de derecha a izquierda, de arriba a abajo de la imagen.

Cada píxel consiste en una (para escala de grises) o tres (si es PFM en color) muestras, una tras otra, sin ningún delimitador entre ellas.

Cada muestra está formada por cuatro bytes consecutivos. Los bytes representan una cadena de 32 bits, tanto en *big* como en *little endian*. Esta cadena de 32 bits se corresponde con un código de número en punto flotante. Como es el mismo formato que la mayoría de las CPUs y compiladores utilizan, se puede programar leyendo los bytes directamente como número en punto flotante (teniendo en cuenta la variación de *endianness*).

Para aclarar posibles dudas, se muestra un fragmento de un fichero en PFM utilizando el visor XVI32 (disponible para descarga gratuita en <http://xvi32.softonic.com>). En

la Figura B.4 se aprecia que el fichero es en color (PF), de tamaño 1920x1080, en *little endian* (-1) y sin escala (*abs*(-1)).

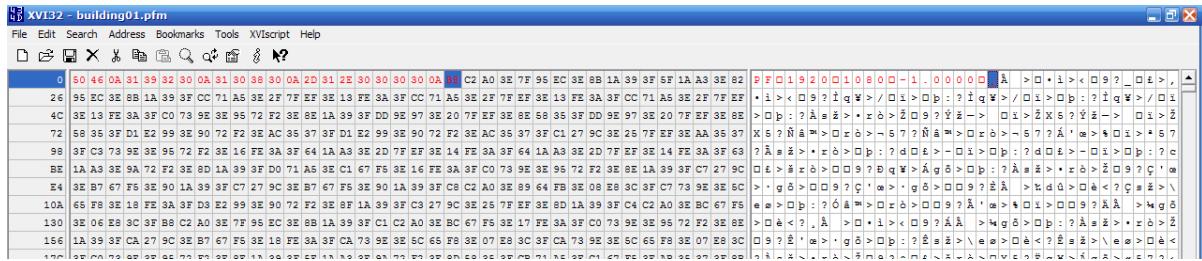


Figura B.4: Gracias a algunos programas como XVI32 podemos visualizar el contenido, byte a byte, de un PFM. En rojo se detalla la cabecera, y en negro parte del cuerpo del PFM.

B.4. Edición

De todas las tareas relacionadas con el proceso de obtención final de una fotografía, la edición es una de las que más se beneficia del alto rango dinámico. Con esta tecnología podemos decidir qué información descartar en el momento de mostrarla y no en el momento del disparo. Esto nos permite contar con una mayor cantidad de información durante el proceso de edición. Gracias a HDR, podemos mejorar el contraste en determinadas zonas, compensar la exposición de la fotografía, etc. Cualquier usuario de Adobe Photoshop [Pho] conoce los problemas que supone la edición con formatos de 8 bits. Si se aclara una imagen la información original se pierde, siendo imposible repararla aunque se oscurezca de nuevo la fotografía. Esto obliga a planear de forma cuidadosa la estrategia de edición. En el futuro, usando información de luminancia real, al aclarar una fotografía y después oscurecerla se obtendrá la imagen original. Trabajar en la edición fotográfica será más rápido y más barato.

La tecnología HDR también permite el diseño y desarrollo de algoritmos mucho más potentes (generación de cáusticas, re-iluminación basada en la imagen, etc) que los permitidos por la tecnología LDR.

Además, surge la posibilidad de utilizar estos mapas de radiancia como fuentes de luz complejas en iluminación de escenas sintéticas. La riqueza visual de la iluminación sería muy complicada de conseguir con los métodos tradicionales de iluminación. Por otro lado, este efecto no se podría realizar con una imagen de bajo rango dinámico, porque la luz emitida desde algunas de las partes de la imagen supera con creces los órdenes de magnitud.

Sin embargo, no todo son ventajas usando HDR. Como cualquier otra tecnología, requiere de un período de habituación para aprovechar su potencial. El principal problema será crear GUI's para edición de HDR. Muchos diseñadores están habituados al sistema

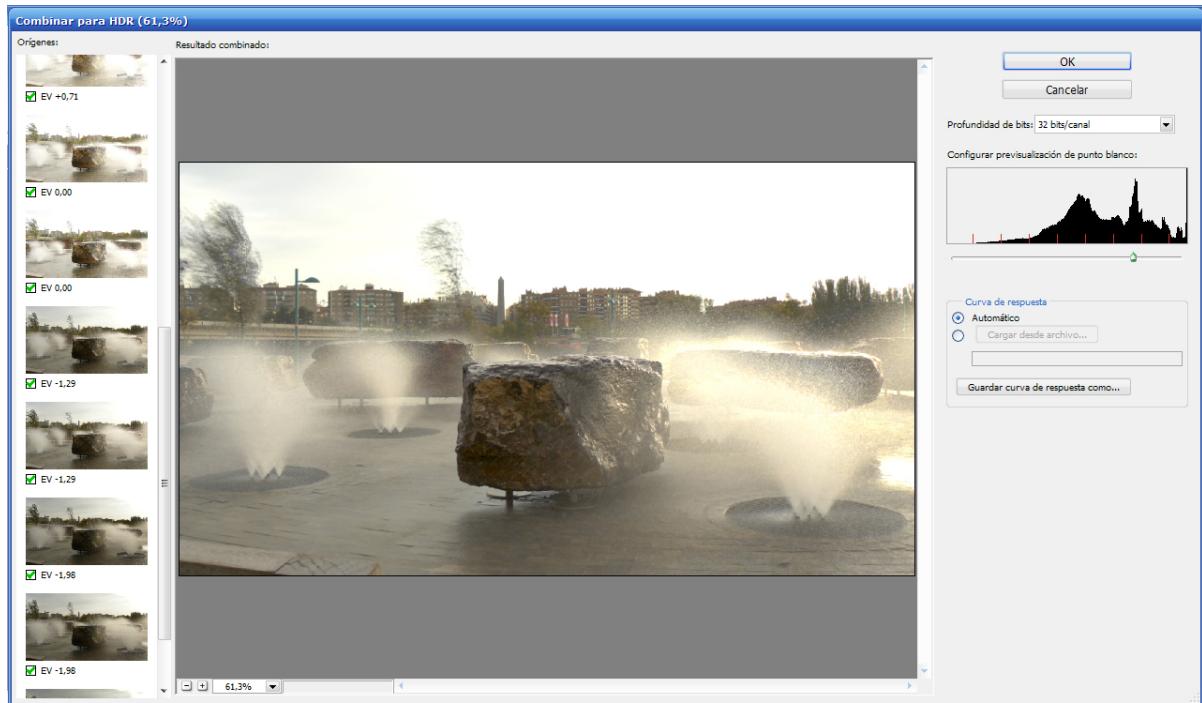


Figura B.5: Adobe Photoshop (combinar para HDR).

RGB-0-255 (en él, la tupla 128,128,128 es un gris medio). Pero, ¿qué sucederá cuando se codifique el color entre 0 y varios billones? ¿Qué código será el gris? ¿Cuál será el blanco? ¿El blanco del monitor? ¿El del papel? ¿La luz del sol?.

Otro inconveniente de la nueva tecnología es que el tamaño de las imágenes se multiplicará por cuatro, al pasar de 8 bits por canal a 32. Esto no sólo implicará problemas de almacenamiento, sino también requerirá de computadores más potentes para la edición (tendrán que manejar cuatro veces más de información).

B.5. Tone Mapping (mapeado de tono)

La mayoría de los monitores (LCD, CRT, TFT) y, por supuesto, las impresoras, tienen un bajo rango dinámico. Así, si se desea mostrar una imagen HDR en papel o en un monitor, se debe convertir de alguna forma el alto rango de la fotografía al aceptado por el medio. A este proceso se le conoce con el nombre de mapeado de tono (o *tone mapping*).

Existen dos tipos de operadores de reproducción de tono: uniformes en el espacio (globales) y variantes en el espacio (locales). Los globales aplican la misma transformación a todos los píxeles, sin tener en cuenta su posición dentro de la imagen. Los operadores locales, por el contrario, aplican diferentes escalas a diferentes zonas de la imagen. Otro aspecto de la reproducción de tono es el tiempo. En [DCWP02b] se hacen las dos distinciones, espacial y temporal, y se realiza una taxonomía de los operadores de tono (Figura

B.6)

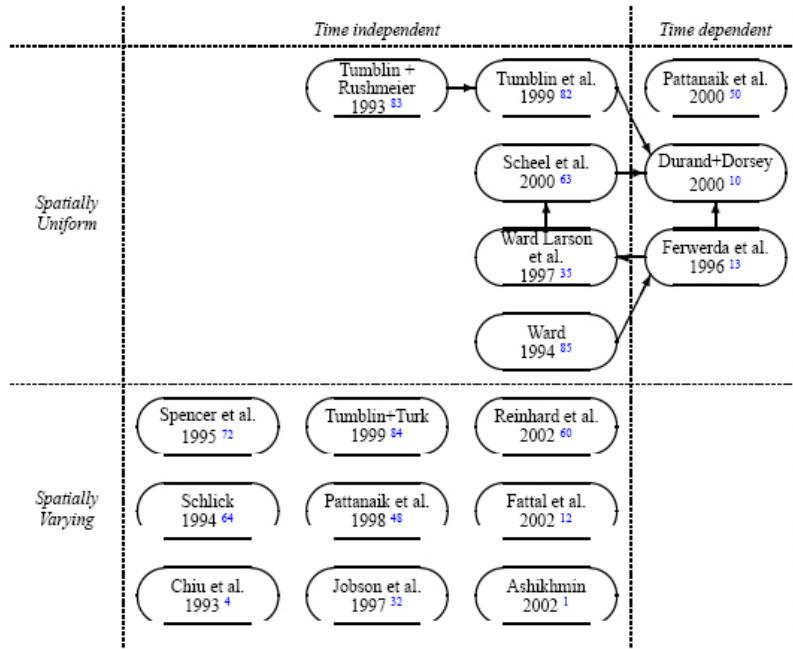


Figura B.6: Taxonomía espacio-temporal de los operadores de tono.

Se han publicado varios operadores de tono, sin embargo ninguno es completamente capaz de generar resultados perceptualmente exactos. Se han realizado diversas comparativas entre ellos (véase [LCTS05]) y es un tema al que se dedican muchos esfuerzos en la actualidad (Según Google Scholar, existen más de 900 artículos de investigación relacionados de forma directa con el proceso del mapeado de tono, que concluyen en varias docenas de algoritmos distintos [MS08]).

Mientras el hardware para visualizar las imágenes HDR no esté del todo desarrollado, el éxito de esta tecnología dependerá en la calidad de los operadores de tono. Ya se han realizado estudios comparativos entre estos operadores, usando monitores convencionales y de alto rango dinámico.

B.6. Visualización

Durante la década pasada, los principales productores de sistemas de visualización intentaron mejorar el rango dinámico de sus monitores. Hoy en día muchos monitores digitales tienen un rango de entorno 2000:1, algo impensable hace años. Esta tendencia va a continuar, sin ninguna duda.

Primero fueron los monitores SVGA CRT (Tubo de Rayos Catódicos) los que propiciaron una auténtica revolución. Les siguieron los monitores LCD y seguidamente los de Plasma. En un futuro inmediato sucederá lo mismo con HDR. Los modelos HDR actuales se basan en una pantalla LCD en la que se ha sustituido la luz trasera, de carácter uniforme en dispositivos convencionales, por una luz cuya intensidad puede variar en el espacio.



Figura B.7: El BrightSide DR37-P es capaz de mostrar un rango dinámico mucho mayor que los monitores actuales.

Los primeros prototipos usaban un proyector con lente Fresnel para iluminar una pantalla LCD [SWW03]. La intensidad del alto rango dinámico era dividida en dos, separando los datos que viajaban a la pantalla LCD de los que viajaban al proyector posterior. Este método presentaba importantes dificultades para alinear de forma correcta el proyector con la pantalla LCD, así como limitaciones en el nivel de negro que podía conseguirse. Estas dificultades hacían inviable la producción comercial del prototipo.

La segunda generación de dispositivos de visualización de alto rango dinámico se ha construido superponiendo una pantalla LCD a una matriz 2D de LEDs de alta luminosidad, que pueden ser modulados de forma independiente [SHS⁺04b]. De esta forma se vencen las dificultades de las versiones basadas en un proyector, pero el coste de la matriz de LEDs, así como las limitaciones de la industria actual, introducen nuevas dificultades, esta vez, de resolución. Por ejemplo, los monitores de 18" se construyen usando 760 LEDs junto con una pantalla LCD de resolución 1280x1024 (Figura B.8). El nivel de negro alcanzado por estos sistemas está alrededor de 0.03 cd/m^2 y su luminancia de pico es 8500 cd/m^2 .

El último modelo es el BrightSide DR37-P (Figura B.7), utilizado en la mayoría de estudios de investigación. Su área de pantalla mide 32.26" por 18.15". Cuenta con una resolución de 1920x1080 píxeles y trabaja con 1380 LEDs. Esta configuración puede dar ratios de contraste que superan 200.000:1, con un nivel de negro de 0.015 cd/m^2 . La luminancia de pico se estima en más de 3000 cd/m^2 .

La Figura B.10 muestra, de forma esquemática, el modo de funcionamiento de este monitor cuando visualizamos una imagen HDR.

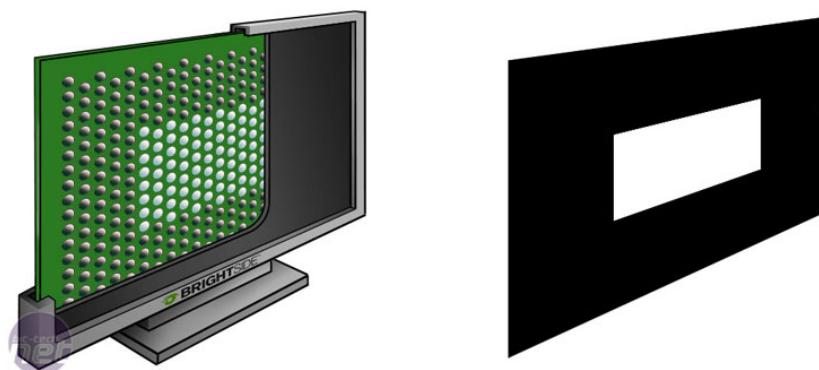


Figura B.8: Los monitores HDR cuentan con una pantalla de LEDs de alta luminosidad superpuesta a una LCD.

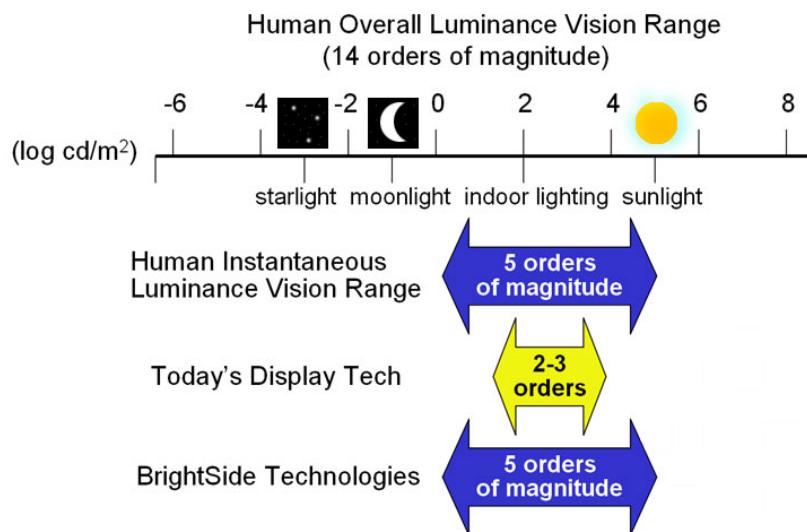


Figura B.9: Rangos lumícos en la naturaleza, mostrados por monitores convencionales y por el BrightSide DR37-P.

El lector podría asumir que el nivel de contraste alcanzado por este dispositivo puede resultar incluso cegador. Lejos de ser así, tras unos minutos viendo imágenes en este monitor, uno se acostumbra rápidamente a la sensación, pues es casi tan natural como mirar la realidad. Basta un vistazo comparativo a la misma imagen, mostrada en este dispositivo y en uno convencional al mismo tiempo, para darnos cuenta de la gran diferencia existente entre ambos en cuanto a calidad y naturalidad de la imagen. Podríamos afirmar que lo único malo de estos monitores ... es su precio.

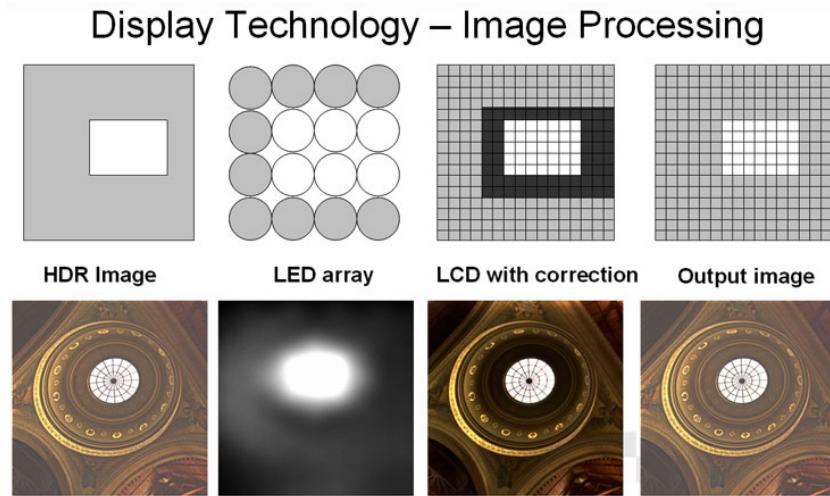


Figura B.10: Funcionamiento de monitores HDR: imagen HDR, iluminación del display de LEDs, corrección e imagen de salida.



Figura B.11: Las diferencias entre el monitor HDR (derecha) y el LDR (izquierda) son muy notables: el blanco es más brillante, y el negro es más oscuro en el HDR. En el LDR se aprecian las consecuencias de la limitación de rango.



Figura B.12: Las diferencias entre el monitor HDR (derecha) y el LDR (izquierda) se aprecian también al observarlos desde un lado.

B.7. Aplicaciones

Hoy en día, el público principal de la tecnología HDR está compuesto por los profesionales del sector de cine, animación y tareas relacionadas con la fotografía, aunque también se le encuentran usos en medicina, geología, física, astronomía. Además de las ventajas enumeradas anteriormente (fundamentalmente en el campo de la fotografía y edición fo-



Figura B.13: A la hora de visualizar fotografías los colores son mucho más naturales en el monitor HDR.

tográfica), encontramos otros curiosos usos derivados de esta corriente tecnológica:

B.7.1. Películas

Herramientas como HDRShop [HDRb], de Paul Debevec, permiten convertir una serie de fotografías en un mapa de luminancias (un tipo de imagen especial que reproduce los valores de iluminación reales de la escena). Estos mapas de luminancia se pueden usar para iluminar objetos virtuales con la luminancia real, de tal forma que parezcan integrados en la escena.

B.7.2. Juegos

Las últimas tarjetas gráficas para ordenadores dan soporte para mapas de texturas HDR. Gracias a ellos se puede iluminar objetos, en tiempo real, usando mapas de luminancia, alcanzando efectos visuales más dinámicos y reales.

B.7.3. Medicina

Gracias a la extrema claridad/oscuridad que puede conseguirse con los nuevos dispositivos de visualización HDR, éstos pueden aproximarse al rendimiento obtenido con una radiografía proyectada sobre una caja de luz. El standard DICOM (*Digital Imaging and Communications in Medicine*) [DIC] para intercambio de fotografías médicas habla de rangos lumínicos desde 0.05 a 4000 cd/m^2 . Monitores como el BrightSide cubren sobradamente este requisito (Figuras B.14 y B.15). Esto implica que las radiografías de los pacientes podrán almacenarse de forma digital, con las consecuentes ventajas que esto acarrea.



Figura B.14: Gracias a la tecnología HDR se podrán almacenar y visualizar digitalmente las radiografías de los pacientes.



Figura B.15: La laparoscopia se beneficiará de la tecnología HDR.

B.7.4. Datos geológicos

En la industria reciente de la extracción de aceites y gases, las compañías invierten millones de dólares al año en localizar reservas no explotadas del “oro negro”. Se realizan sofisticados análisis geofísicos de fondos marinos (Figura B.16), que producen grandes cantidades de datos que deben ser visualizados y analizados (ardua tarea cuando los monitores se limitan a 8 bits por canal).

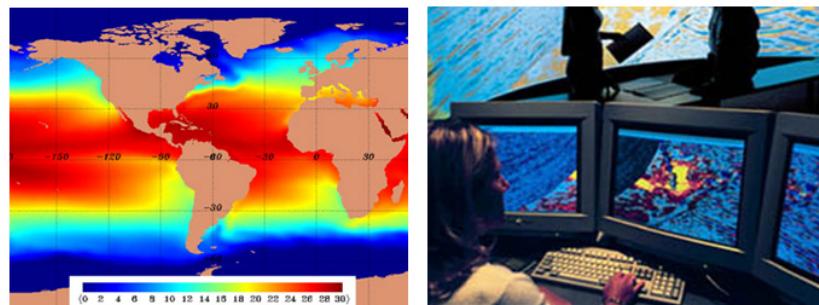


Figura B.16: HDR contribuirá a facilitar la labor de análisis geofísicos, porque enriquecerá el proceso de visualización de datos.

B.7.5. Imágenes por satélite

Los satélites de hoy día capturan datos en 16 bits por canales, pese a que todos los monitores existentes son únicamente de 8 bits. En el futuro, científicos y astrónomos podrán estudiar el rango dinámico completo de la imagen original.

B.7.6. Nuevos algoritmos

La tecnología HDR ha permitido el desarrollo y aplicación de nuevos algoritmos (por ejemplo, generación de imágenes foto-realistas mediante iluminación basada en la imagen [Bla06]), que habrían sido imposibles de desarrollar con la tecnología LDR.

Apéndice C

Psicofísica

C.1. Psicofísica: conceptos

El concepto de psicofísica aparece cuando Gustav Theodor Fechner (1801-1887) introduce en 1860 el nuevo programa de investigación interdisciplinar, llamado *Psychophysik* [GT], que presenta un método científico para estudiar las relaciones entre el cuerpo y la mente. El modelo matemático que presentó se conoce como la Ley de Weber-Fechner. ésta establece una relación cuantitativa entre la magnitud de un estímulo físico y como este es percibido. Fue propuesta en primer lugar por Ernst Heinrich Weber (1795-1878), y elaborada hasta su forma actual por Fechner. Ernst Heinrich Weber estableció su ley de la sensación (o Ley de Weber) en la que formulaba la relación matemática que existía entre la intensidad de un estímulo y la sensación producida por éste. Estos y otros descubrimientos llevaron a la convicción de que era posible explicar mediante principios físico-químicos todos los actos humanos. De este modo Fechner se anticipaba a uno de los objetivos principales de la neurología moderna [EE] .

Según Wikipedia [Wikc] , la psicofísica es *una rama de la psicología que estudia la relación entre la magnitud de un estímulo físico y la intensidad con la que éste es percibido por parte de un observador. Su objetivo es poder hallar un escalamiento en el que situar esta relación.*

John C. Baird y Elliot Noma ofrecen una definición completa del concepto de psicofísica en la primera página de su libro [JWS78] :

La psicofísica se define comúnmente como la rama cuantitativa del estudio sobre la percepción, examinando las relaciones entre los estímulos observados y las respuestas y razones para estas relaciones [...] Desde su concepción, la psicofísica se ha basado en asumir que el sistema perceptual humano es un instrumento de medida que produce resultados (experiencias, juicios y respuestas) que pueden ser analizadas sistemáticamente. Debido a su experiencia (más de 140 años), sus métodos experimentales, análisis de datos y modelos de percepción y cognitivos han alcanzado un alto nivel de refinamiento. Por este motivo, muchas técnicas originalmente desarrolladas en psicofísica se han utilizado para desentrañar problemas de aprendizaje, memoria y psicología social. Además, en muchos ámbitos

de toma de decisiones se han adaptado las teorías de la medición y el escalado de sistemas.

Esta rama de la psicología, por tanto, tiene como fin medir objetivamente lo subjetivo, por medio de umbrales. La psicofísica está íntimamente relacionada con la función psíquica que permite al organismo, a través de los sentidos, recibir, elaborar e interpretar la información proveniente de su entorno; es decir, la percepción.

C.2. Prácticas empleadas

En el proyecto que nos ocupa se emplearán varios tests psicofísicos. Para todos ellos se utilizarán fotografías (cuidadosamente elegidas) como estímulos. La tarea de los encuestados será asignar una puntuación a cada fotografía, para cada uno de los aspectos en que sean interrogados. Nótese que estas preguntas serán idénticas, dentro del mismo test, para los distintos estímulos.

Todos los tests realizados se desarrollarán según una metodología cuidada, que respeta las buenas prácticas psicofísicas. Las más destacables son:

C.2.1. Periodo de *entrenamiento*

Todo test contará con un periodo de *entrenamiento*, en el que los participantes tomen experiencia, sin que sus resultados en esta fase sean tenidos en cuenta. Sin embargo, los participantes no serán conocedores de ésto. Así se consigue que el participante cuente con algo de *criterio*, en base a lo aprendido en el periodo de entrenamiento, para velar por la integridad de las respuestas obtenidas.

En el caso que nos ocupa, los participantes observarán cinco fotografías hasta que sus respuestas se tengan en cuenta. A partir de ese momento las respuestas de los encuestados serán computadas para obtener conclusiones.

C.2.2. Aleatorización de muestras

Los estímulos serán aleatorizados deliberadamente, de un encuestado a otro. De esta manera se tratará de suavizar el efecto de la experiencia adquirida según se va avanzando en la realización del test.

En nuestro caso las imágenes se mostrarán el orden aleatorio. Si siempre se mostrasen en el mismo orden, podría ocurrir que la comparación entre una imagen y la siguiente afectara a los resultados, incluso aunque el encuestado no lo advirtiera.

C.2.3. Escala de medición conocida

La escala de valoración empleada debe resultar familiar a los encuestados. Si están habituados a su uso podrán emplearla de forma más cómoda. En cada país o región es costumbre utilizar una escala u otra.

En el caso que nos ocupa se utilizará una escala muy simple: El valor '1' es el mejor, cualitativamente hablando. Le siguen los valores '2', '3' y '4', de mejor a peor.

<i>Puntuación</i>	<i>Presencia de estímulo</i>
0	Ninguna en absoluto.
1	Extremadamente débil (casi imperceptible).
2	Muy débil.
3	Débil (ligera).
4	Moderada.
5	Presente.
6	Notable.
7	Muy notable.
8	Fuerte.
9	Muy fuerte.
10	Total.

Tabla C.1: Escala BorgCR10.

No se puede esperar que todo el mundo sea capaz de clasificar de forma fiable según ninguna escala existente. Un pequeño porcentaje de los adultos (alrededor del cinco por cien) tiene dificultades para entender las instrucciones y peticiones utilizando la escala Borg CR10 (Figura C.1). Este fenómeno aumenta para individuos con poca habilidad verbal y matemática [Bor].

C.2.4. Pùblico heterogéneo

El propósito de este proyecto es que las imágenes legadas se muestren correctamente en los monitores de nueva generación. El público potencialmente interesado, y consumidor en un futuro inmediato de esta tecnología es muy amplio, abarcando distintas edades, sexos, culturas, niveles económicos y educativos. Es por ello que se tratará, en la medida de lo posible, de contar con un conjunto heterogéneo de encuestados de cara a la realización de los tests. No interesa únicamente la opinión de ingenieros, fotógrafos o técnicos. Interesa contar con una muestra representativa de la sociedad.

C.2.5. Elementos de control

De cara a asegurar la integridad de los datos recopilados en los tests, suelen utilizarse dos técnicas. La primera repetir el test completo a cada individuo, realizando las preguntas en otro orden. La segunda es introducir algunos estímulos repetidos, a fin de comprobar que las respuestas obtenidas por el encuestado son íntegras. El porcentaje de estímulos repetidos frente a los originales varía en función del test realizado.

Nótese que, pese a ser la primera técnica más precisa, es muy poco común en tests de larga duración, pues repetir el test puede resultar pesado para el encuestado, al que los estímulos y preguntas le resultan familiares. ésto suele derivar en falta de atención y resultados imprecisos.

En el caso que nos ocupa, tomaremos la siguiente regla: si la duración estimada del test supera los veinte minutos se introducirán un diez por ciento de imágenes de control. Si no supera los veinte minutos, se repetirá el test al encuestado (no de forma inmediata a la realización del primero, sino dejando pasar unas horas).

C.2.6. Libertad de modificación

Se asume que todo encuestado irá adquiriendo experiencia en la contestación de preguntas al tiempo que realiza el test. Por tanto, es probable que en determinado momento de la prueba, al observar con detenimiento cierto estímulo, el individuo considere oportunuo modificar la valoración de alguno de los estímulos anteriores.

En los tests realizados en este proyecto se permitirá la libre modificación de respuestas anteriores, permitiendo al usuario visualizar de nuevo cuantos estímulos mostrados en el pasado desee, de forma que pueda realizar cuantas modificaciones quiera.

Apéndice D

Estímulos

En este Apéndice se mostrarán las imágenes utilizadas para los tests psicofísicos. Se ha buscado una amplia variedad de iluminaciones (luz natural: amaneceres, atardeceres, día soleado, día nublado y luz artificial), situaciones (interiores, exteriores) y motivos (paisajes, objetos cercanos y lejanos, etc).

Debajo de cada imagen se muestra el histograma de la misma. Nótese que al aumentar la exposición de las secuencias los valores del histograma se desplazan hacia la derecha.

Además, para cada imagen se muestran los datos estadísticos (Media¹, Desv.Típ.², Mediana³, porcentaje de píxeles con luminancia 0 y con luminancia 255 respecto al total de píxeles.) asociados a su histograma más relevantes, siguiendo el formato de la Tabla D.1

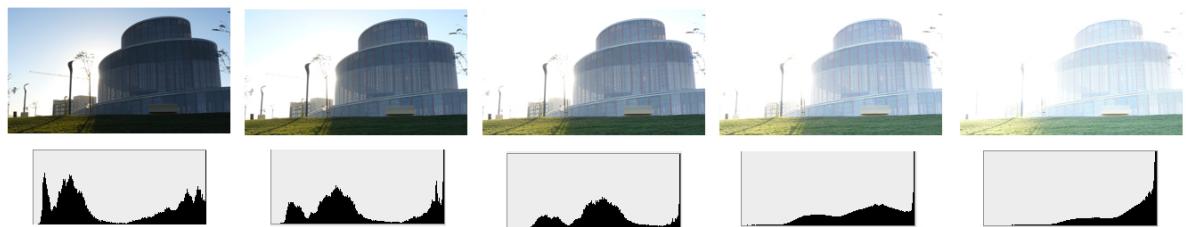
Media(img ₁)	Media(img ₂)	Media(img ₃)	Media(img ₄)	Media(img ₅)
Desv.Típ.(img ₁)	Desv.Típ.(img ₂)	Desv.Típ.(img ₃)	Desv.Típ.(img ₄)	Desv.Típ.(img ₅)
Mediana(img ₁)	Mediana(img ₂)	Mediana(img ₃)	Mediana(img ₄)	Mediana(img ₅)
%(img ₁ ,0)	%(img ₂ ,0)	%(img ₃ ,0)	%(img ₄ ,0)	%(img ₅ ,0)
%(img ₁ ,255)	%(img ₂ ,255)	%(img ₃ ,255)	%(img ₄ ,255)	%(img ₅ ,255)

Tabla D.1: Formato de datos estadísticos de los estímulos.

¹ La media aritmética (o promedio) de una cantidad finita de números es igual a la suma de todos ellos dividida entre el número de sumandos. $\bar{x} = \frac{1}{n} \cdot \sum_{i=1}^n x_i$

² La desviación estándar (desviación típica) es una medida (cuadrática) que informa de la media de distancias que tienen los datos respecto de su media aritmética, expresada en las mismas unidades que la variable. $\sqrt{s^2} = \sqrt{\frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n-1}}$

³ Mediana es el valor de la variable que deja el mismo número de datos antes y después que él. Considerando $x_1, x_2, x_3, \dots, x_n$ los datos de una muestra ordenada en orden creciente y mediana como M_e : $M_e = x_{\frac{(n+1)}{2}}$, si n es impar $\rightarrow M_e$ será la observación central de los valores, una vez que estos han sido ordenados en orden creciente o decreciente. $M_e = \frac{x_{\frac{n}{2}} + x_{\frac{n}{2}+1}}{2}$, si n es par $\rightarrow M_e$ será el promedio aritmético de las dos observaciones centrales.

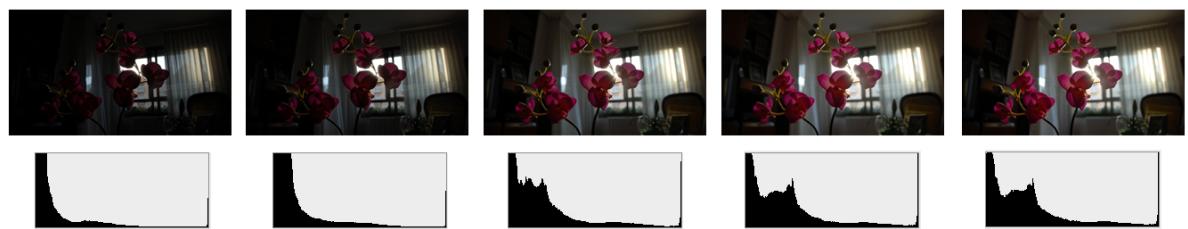
Figura D.1: Secuencia 01: *building*. Imagenes e histogramas.

126.66	158.96	183.77	211.39	232.45
90.02	86.2	72.99	54.26	37.30
85	137	189	240	255
0 %	0 %	0 %	0 %	0 %
4.39 %	18.78 %	38.93 %	43.82 %	46.66 %

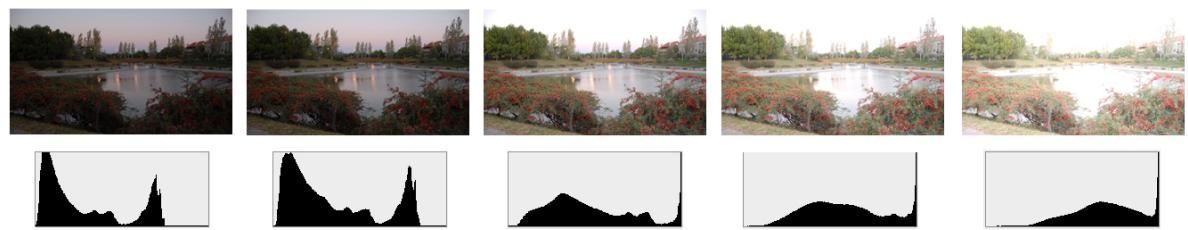
Tabla D.2: Datos estadísticos de los estímulos secuencia *building*Figura D.2: Secuencia 02: *car*. Imagenes e histogramas.

14.72	23.79	38.41	58.58	85.06
29.92	38.84	50.68	62.40	73.07
6	14	27	50	86
30.66 %	19.65 %	9.23 %	4.08 %	1.16 %
0.01 %	0.01 %	0.18 %	0.56 %	1.84 %

Tabla D.3: Datos estadísticos de los estímulos secuencia *car*

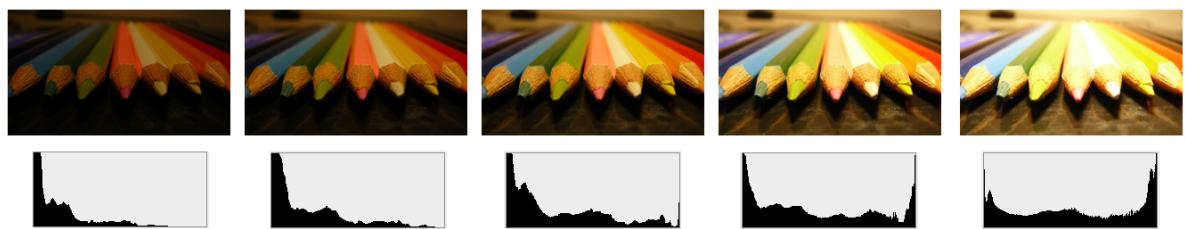
Figura D.3: Secuencia 03: *indoor*. Imágenes e histogramas.

19.67	30.29	50.28	60.41	60.51
37.23	45.22	56.82	61.63	61.79
7	15	33	45	45
26.11 %	17.10 %	5.99 %	3.62 %	3.50 %
0.58 %	0.73 %	1.21 %	1.85 %	1.87 %

Tabla D.4: Datos estadísticos de los estímulos secuencia *indoor*Figura D.4: Secuencia 04: *lake*. Imágenes e histogramas.

70.77	86.47	138.04	170.17	201.54
62.19	68.16	76.33	67.19	53.52
43	61	119	164	210
0.10 %	0 %	0 %	0 %	0 %
0 %	0 %	17.91 %	22.79 %	33.35 %

Tabla D.5: Datos estadísticos de los estímulos secuencia *lake*

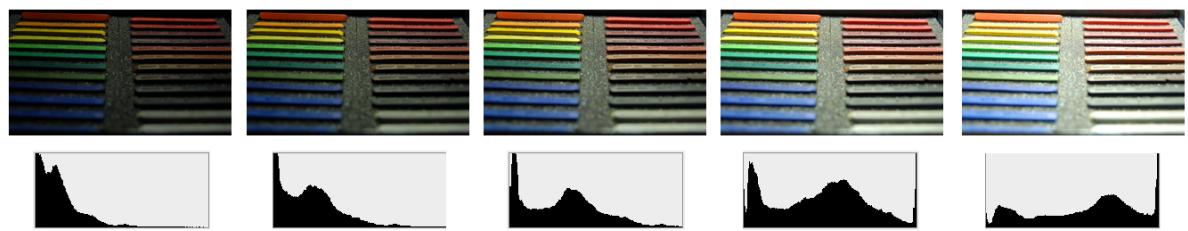
Figura D.5: Secuencia 05: *pencils*. Imágenes e histogramas.

26.68	43.37	68.65	104.5	138.95
37.47	53.35	70.10	83.16	87.10
9	18	39	83	136
12.51 %	8.38 %	5.04 %	1.02 %	0.06 %
0 %	0 %	0 %	0 %	1.94 %

Tabla D.6: Datos estadísticos de los estímulos secuencia *pencils*Figura D.6: Secuencia 06: *computers*. Imágenes e histogramas.

14.93	20.55	24.05	34.98	52.51
48.15	56.18	59.04	65.32	67.64
0	3	5	11	26
58.06 %	23.53 %	14.11 %	2.70 %	0.12 %
0.09 %	0.43 %	0.72 %	2.07 %	7.18 %

Tabla D.7: Datos estadísticos de los estímulos secuencia *computers*

Figura D.7: Secuencia 07: *waxes*. Imágenes e histogramas.

29.48	50.41	78.21	111.85	148.84
30.02	43.00	56.37	67.58	73.99
23	46	80	120	167
13.13 %	4.56 %	0.45 %	0 %	0 %
0 %	0 %	0 %	0 %	2.11 %

Tabla D.8: Datos estadísticos de los estímulos secuencia *waxes*Figura D.8: Secuencia 08: *sunset*. Imágenes e histogramas.

91.86	108.56	139.87	194.62	229.99
73.20	74.18	78.10	68.39	46.36
75	97	136	218	255
0 %	0 %	0 %	0 %	0 %
3.53 %	4.43 %	8.30 %	26.01 %	51.48 %

Tabla D.9: Datos estadísticos de los estímulos secuencia *sunset*

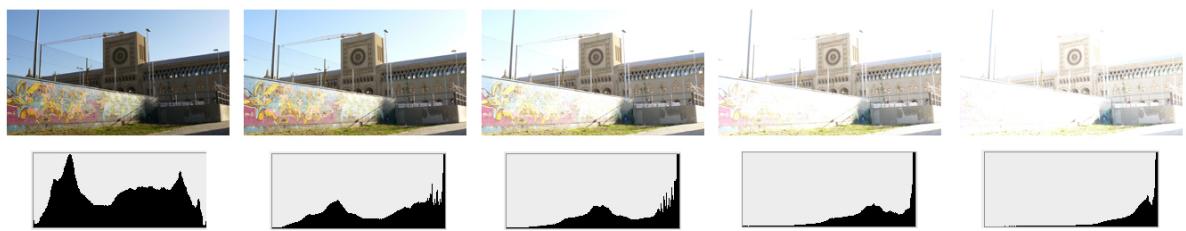


Figura D.9: Secuencia 09: *graffitti*. Imágenes e histogramas.

127.06	169.69	202.56	226.57	242.94
69.86	71.90	59.51	41.48	25.24
133	189	233	255	255
0 %	0 %	0 %	0 %	0 %
0 %	0.58 %	17.72 %	47.77 %	59.84 %

Tabla D.10: Datos estadísticos de los estímulos secuencia *graffitti*

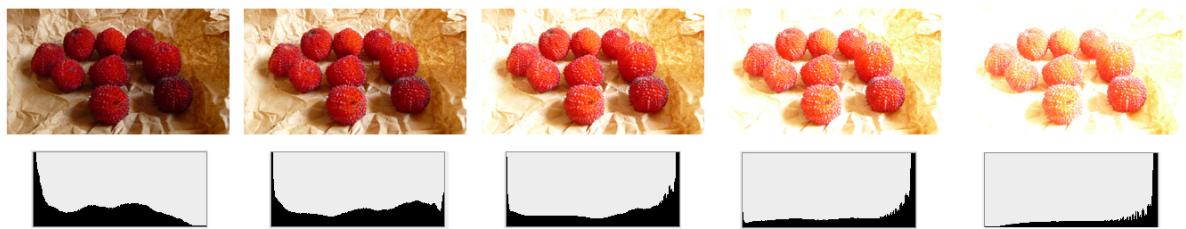


Figura D.10: Secuencia 08: *strawberries*. Imágenes e histogramas.

90.39	130.18	168.60	199.79	223.45
67.35	82.11	85.22	75.27	57.46
90	140	198	246	254
0 %	0 %	0 %	0 %	0 %
0 %	0 %	5.84 %	21.94 %	37.94 %

Tabla D.11: Datos estadísticos de los estímulos secuencia *strawberries*

Apéndice E

Artículo de investigación presentado al CEIG2008

En este Apéndice se adjunta el artículo de investigación presentado al CEIG2008, que sintetiza muchos de los aspectos estudiados y trabajados en este estudio y cuyo autor principal es el mismo que el de este Proyecto.

Bibliografía

- [AFR⁺07] Ahmet Oğuz Akyüz, Roland Fleming, Bernhard E. Riecke, Erik Reinhard, and Heinrich H. Bülthoff. Do hdr displays support ldr content?: a psychophysical evaluation. In *SIGGRAPH '07: ACM SIGGRAPH 2007 papers*, page 38, New York, NY, USA, 2007. ACM.
- [Aky] Ahmet Oguz Akyüz. Homesite (http://www.coolhall.com/homepage/pubs/hdrdisp_eval/hdrdisp_project.html).
- [BCJC] F. Bartz, D. Cunningham, J.F., and Wallraven C. The role of perception for computer graphics. In *State of the Art Reports, Eurographics 2008*.
- [Bla06] Francisco Meléndez Blas. *Generación de imágenes foto-realistas mediante iluminación basada en la imagen*. 2006.
- [BLDC06] Francesco Banterle, Patrick Ledda, Kurt Debattista, and Alan Chalmers. Inverse tone mapping. In *GRAPHITE '06: Proceedings of the 4th international conference on Computer graphics and interactive techniques in Australasia and Southeast Asia*, pages 349–356, New York, NY, USA, 2006. ACM.
- [BLLC02] Raphaél Bornard, Emmanuelle Lecan, Louis Laborelli, and Jean-Hugues Chenot. Missing data correction in still images and image sequences. In *ACM Multimedia*, pages 355–361, 2002.
- [Bor] G. Borg.
- [BS08] Mario Botsch and Olga Sorkine. On linear variational surface deformation methods. *IEEE Transactions on Visualization and Computer Graphics*, 14(1):213–230, 2008.
- [BSS07] Tamy Boubekeur, Olga Sorkine, and Christophe Schlick. SIMOD: Making freeform deformation size-insensitive. In *Proceedings of IEEE/Eurographics Symposium on Point-Based Graphics*, pages 47–56, 2007.
- [BVS03] M. Bertalmio, L. Vese, G. Sapiro, and S. Osher. Simultaneous structure and texture image inpainting. *cvpr*, 02:707, 2003.
- [Coh88] J. Cohen. *Statistical power analysis for the behavioral sciences*. Hillsdale, NJ: Lawrence Erlbaum Associates, 1988.

- [Com90] International Radio Consultative Committee(ITU). Basic parameter values for the hdtv standard for the studio and for international programme exchange, 1990. url <http://www.w3.org/Graphics/Color/sRGB>.
- [DCWP02a] Kate Devlin, Alan Chalmers, Alexander Wilkie, and Werner Purgathofer. Star: Tone reproduction and physically based spectral rendering. In Dieter Fellner and Roberto Scopigno, editors, *State of the Art Reports, Eurographics 2002*, pages 101–123. The Eurographics Association, September 2002.
- [DCWP02b] Kate Devlin, Alan Chalmers, Alexander Wilkie, and Werner Purgathofer. STAR: Tone Reproduction and Physically Based Spectral Rendering. In Dieter Fellner and Roberto Scopigno, editors, *State of the Art Reports, Eurographics 2002*, pages 101–123. The Eurographics Association, September 2002.
- [Deb] Paul Debevec homesite. url <http://www.debevec.org/>.
- [Deb05] Paul Debevec. A median cut algorithm for light probe sampling. In *SIGGRAPH '05: ACM SIGGRAPH 2005 Posters*, page 66, New York, NY, USA, 2005. ACM.
- [DIC] Estándar DICOM. url <http://es.wikipedia.org/wiki/DICOM>.
- [DM97] Paul E. Debevec and Jitendra Malik. Recovering high dynamic range radiance maps from photographs. *Computer Graphics*, 31(Annual Conference Series):369–378, 1997.
- [DP-] DP-review. Review of Fuji's SR chips. url <http://www.dpreview.com/news/0301/03012202fujisuperccdsr.asp>.
- [EE] Walter H. Ehrenstein and Addie Ehrenstein. Psychophysical methods. url <http://darwin.psy.ulaval.ca/isp/history/texts/PSYPHY-M.PDF>.
- [ElG] El Greco, “La agonía en el jardín”. url <http://www.ceciliabalza.com.ar/obras/elgreco.html>.
- [ESA07] Mathias Eitz, Olga Sorkine, and Marc Alexa. Sketch based image deformation. In *Proceedings of Vision, Modeling and Visualization (VMV)*, pages 135–142, 2007.
- [FABJ04] R. W. Fleming, E. H. Adelson, H.H. Bülthoff, and H. W. Jensen. Perceiving translucent materials, May 2004.
- [FTA04] R. W. Fleming, A. Torralba, and E. H. Adelson. Specular reflections and the perception of shape. *Journal of Vision*, 4(9):798–820, 2004.
- [Gam] Gamma FAQ. url <http://www.poynton.com/GammaFAQ.html>.
- [GIF] Formato GIF. url http://es.wikipedia.org/wiki/Graphics_Interchange_Format.

- [GKB⁺99] Alan Gilchrist, Christos Kossyfidis, Frederick Bonato, Tiziano Agostini, Joseph Cataliotti, Xiaojun Li, Branka Spehar, Vidal Annan, and Elias Economou. An anchoring theory of lightness perception. *Psychological Review*, 106(4):795–834, October 1999.
- [Gor] The Prokudin Gorskii Photographic Record. url <http://www.loc.gov/exhibits/empire/>.
- [GT] Fechner GT. *Elemente der Psychophysik*. Breitkopf & Härtel, Leipzig (reprinted in 1964 by Bonset, Amsterdam); English translation by HE Adler (1966): Elements of psychophysics. Holt, Rinehart & Winston, New York.
- [HDRa] Brightside HDR display review by *bit-tech.net*. url http://www.bit-tech.net/hardware/2005/10/04/brightside_hdr_edr.
- [HDRb] HDRshop. HDRshop website. url <http://www.hdrshop.com> .
- [HT96] Anil N. Hirani and Takashi Totsuka. Combining frequency and spatial domain information for fast interactive image noise removal. *Computer Graphics*, 30(Annual Conference Series):269–276, 1996.
- [ITU] Itu. 1990. international telecommunication union itu-r recommendation bt.709, basic parameter values for the hdtv standard for the studio and for international programme exchange. geneva. formerly ccir rec. 709.
- [JLR07] JLR. 8 consejos clave para hacer fotos hdr espectaculares, 2007. url <http://www.dzoom.org.es/noticia-1584.html>.
- [JPEa] Formato JPEG. url <http://en.wikipedia.org/wiki/JPEG>.
- [JPEb] Formato JPEG2000. url <http://www.jpeg.org/jpeg2000/>.
- [JWS78] Inc John Wiley & Sons. *Fundamentals of Scaling and Psychophysics*. 1978.
- [Kar] Karol myszkowski website. url <http://www.mpi-inf.mpg.de/~karol/>.
- [Ken75] M. Kendall. *Rank Correlation Methods*. Charles Griffin & Co. Ltd, 1975.
- [Kor] N. Koren. Tonal quality and dynamic range in digital cameras. url http://www.normankoren.com/digital_tonality.html#Human_vision.
- [LCTS05] Patrick Ledda, Alan Chalmers, Tom Troscianko, and Helge Seetzen. Evaluation of tone mapping operators using a High Dynamic Range display. *ACM Trans. Graph.*, 24(3):640–648, 2005.
- [Liv02] M. Livingstone. *Vision and Art: The Biology of Seeing*. Harry N. Abrams., 2002.
- [LLW04] Anat Levin, Dani Lischinski, and Yair Weiss. Colorization using optimization. *ACM Trans. Graph.*, 23(3):689–694, August 2004.

- [MDS06] Laurence Meylan, Scott Daly, and Sabine S \tilde{A} sstrunk. The Reproduction of Specular Highlights on High Dynamic Range Displays. In *IS&T/SID 14th Color Imaging Conference*, 2006.
- [MDS07] Laurence Meylan, Scott Daly, and Sabine S \tilde{A} sstrunk. Tone Mapping For High Dynamic Range Displays. In *Electronic Imaging*, volume 6492, 2007.
- [Mey] J. Meyer. The future of digital imaging. url <http://www.cybergrain.com/tech/hdr/>.
- [MFSG08] Miguel Martin, Roland Fleming, Olga Sorkine, and Diego Gutierrez. Understanding exposure for reverse tone mapping. 2008.
- [MPIa] Max Planck Institut für Biologische kybernetik. url <http://www.kyb.mpg.de/>.
- [MPIb] Max Planck Institut für Informatik. url <http://www.mpi-inf.mpg.de/>.
- [MS08] Rafa? Mantiuk and Hans-Peter Seidel. Modeling a generic tone-mapping operator. *Computer Graphics Forum*, 27(2):699–708, 2008.
- [Ope] Formato OpenEXR. url <http://www.openexr.com/>.
- [PFM] Formato PFM. url <http://netpbm.sourceforge.net/doc/pfm.html>.
- [Pho] Adobe Photoshop Website. url <http://www.adobe.com/es/products/photoshop/photoshop.html>.
- [PNG] Formato PNG. url <http://www.libpng.org/pub/png/>.
- [RAE] RAE. Real Academia de la Lengua. url <http://www.rae.es/>.
- [RAW] Formato RAW. url http://es.wikipedia.org/wiki/RAW_%28formato%29.
- [Rol] Roland fleming website. url <http://www.kyb.mpg.de/roland>.
- [rsf] rsfotodigital.net. Curso de fotografía de rsfotodigital.net.
- [RSSF02] E. Reinhard, M. Stark, P. Shirley, and J. Ferwerda. Photographic tone reproduction for digital images, 2002.
- [RTS⁺07] Allan G. Rempel, Matthew Trentacoste, Helge Seetzen, H. David Young, Wolfgang Heidrich, Lorne Whitehead, and Greg Ward. Ldr2hdr: on-the-fly reverse tone mapping of legacy video and photographs. In *SIGGRAPH '07: ACM SIGGRAPH 2007 papers*, page 39, New York, NY, USA, 2007. ACM.
- [RWPD05] Erik Reinhard, Greg Ward, Sumanta Pattanaik, and Paul Debevec. *High Dynamic Range Imaging: Acquisition, Display and Image-Based Lighting*. Morgan Kaufmann Publishers, 2005.
- [SAM96] M. Stokes, M. Anderson, and R. Motta. A standard default color space for the internet, 1996. url <http://www.w3.org/Graphics/Color/sRGB>.

- [SH06] Li H. Ye L. Ward G. Whitehead L. Seetzen, H. and W. Heidrich. Guidelines for contrast, brightness, and amplitude resolution of displays. In *In Society for Information Display (SID) Digest*, pages 1229–1233, 2006.
- [SHS⁺04a] H. Seetzen, W. Heidrich, W. Stuerzlinger, G. Ward, L. Whitehead, M. Trentacoste, A. Ghosh, and A. Vorozcova. High dynamic range display systems. *Proceedings of ACM Transactions on Graphics*, 23(3):760–768, 2004.
- [SHS⁺04b] H. Seetzen, W. Heidrich, W. Stuerzlinger, G. Ward, L. Whitehead, M. Trentacoste, A. Ghosh, and A. Vorozcova. High dynamic range display systems. *Proceedings of ACM Transactions on Graphics*, 23(3):760–768, 2004.
- [Sun08] Veronica Sundstedt. A psychophysical study of foveal gradient based selective rendering. In *Proceedings of GRAPP 2008 - Third International Conference on Computer Graphics Theory and Applications*, pages 207–214. INSTICC Press, January 2008.
- [SWW03] Helge Seetzen, Lorne A. Whitehead, and Greg Ward. P. 54.2: A high dynamic range display using low and high resolution modulators, 2003. [url](http://citeseer.ist.psu.edu/709820.html)
- [TM98] C. Tomasi and R. Manduchi. Bilateral filtering for gray and color images. In *ICCV '98: Proceedings of the Sixth International Conference on Computer Vision*, page 839, Washington, DC, USA, 1998. IEEE Computer Society.
- [Vap95] Vladimir N. Vapnik. *The nature of statistical learning theory*. Springer-Verlag New York, Inc., New York, NY, USA, 1995.
- [WAM02] Tomohisa Welsh, Michael Ashikhmin, and Klaus Mueller. Transferring color to greyscale images. In *SIGGRAPH '02: Proceedings of the 29th annual conference on Computer graphics and interactive techniques*, pages 277–280, New York, NY, USA, 2002. ACM Press.
- [Wan95] B. A. Wandell. *Foundations of Vision*. Sinauer Associates, Inc., 1995.
- [Web] Test Accesibilidad Web. url: (<http://www.tawdis.net/taw3/cms/es>).
- [Wika] Wikipedia. Concepto de Apertura. url
<http://es.wikipedia.org/wiki/Apertura>.
- [Wikb] Wikipedia. Concepto de Profundidad de Campo. url
http://es.wikipedia.org/wiki/Profundidad_de_campo.
- [Wikc] Wikipedia. Concepto de Psicofísica. url
<http://es.wikipedia.org/wiki/Psicofísica>.

- [Wikd] Wikipedia. Concepto de *Endianness*. url
<http://en.wikipedia.org/wiki/Endianness>.
- [Wike] Wikipedia. Concepto de Velocidad de Obturación. url
[http://es.wikipedia.org/wiki/Velocidad_de_obturacion_\(fotografia\)](http://es.wikipedia.org/wiki/Velocidad_de_obturacion_(fotografia))
- .

Understanding exposure for reverse tone mapping

Miguel Martin¹, Roland Fleming², Olga Sorkine³ and Diego Gutierrez¹

¹Universidad de Zaragoza, Spain

²Max Planck Institute for Biological Cybernetics, Germany

³New York University, USA

Abstract

High dynamic range (HDR) displays are capable of providing a rich visual experience by boosting both luminance and contrast beyond what conventional displays can offer. We envision that HDR capture and display hardware will soon reach the mass market and become mainstream in most fields, from entertainment to scientific visualization. This will necessarily lead to an extensive redesign of the imaging pipeline. However, a vast amount of legacy content is available, captured and stored using the traditional, low dynamic range (LDR) pipeline. The immediate question that arises is: will our current LDR digital material be properly visualized on an HDR display? The answer to this question involves the process known as reverse tone mapping (the expansion of luminance and contrast to match those of the HDR display) for which no definite solution exists.

This paper studies the specific problem of reverse tone mapping for imperfect legacy still images, where some regions are under- or overexposed. First, we show the results of a psychophysical study compared with first-order image statistics, in an attempt to gain some understanding in what makes an image be perceived as incorrectly exposed; second, we propose a methodology to evaluate existing reverse tone mapping algorithms in the case of imperfect legacy content.

Categories and Subject Descriptors (according to ACM CCS): I.4.0 [Image Processing and Computer Vision]: General–Image Displays I.3.3 [Computer Graphics]: Picture/Image Generation H.1.2 [Models and Principles]: User/Machine Systems Human factors–Human Information Processing

1. Introduction

High dynamic range imagery allows a broad range of physically-accurate photometric values to be stored per pixel, mimicking the ranges that can be perceived by the human visual system [RWP05]. The well-known process of tone mapping [DCWP02] deals with the problem of strong contrast reduction of the stored HDR radiance values to fit the low dynamic range of traditional display technology, typically trying to preserve image details and/or color appearance.

The problem of tone mapping is expected to progressively fade away when HDR displays reach the mass market [SHS*04]. However, during the logical transition period, there will be a need to display conventional low dynamic range (LDR) imagery on HDR displays. Although this need may decline over time (once HDR capture becomes mainstream), 8-bit photography will most likely still be used for a long time. This means that display algorithms will have

to scale up luminance and contrast, instead of compressing them. This brings about the problem of *reverse tone mapping*[†], to which currently no definite solution exists. Recently, Seetzen et al. [SH06] and Yoshida et al. [YMMS06] showed that the subjective perception of image quality increases when both brightness and contrast are increased simultaneously. This result indicates that merely emulating LDR characteristics on an HDR display is probably not the best option, as suggested in [RTS*07].

Very few works exist that deal with the problem of reverse tone mapping. Banterle and colleagues [BLDC06, BLD*07]

[†] Some authors [BLDC06, AFR*07] refer to the process as *inverse* tone mapping, while others [RTS*07] use the term *reverse* instead. Given that the field is still in its infancy, a fixed nomenclature has not been chosen yet. We opt to use *reverse* since the term *inverse* can also refer specifically to mathematically inverting a tone mapping operator, not to the whole process.

propose a method by first inverting Reinhard's tone mapping operator [RSSF02]. The authors then find areas of high luminance and apply density estimation techniques to produce an *expand-map*, which guides the range expansion of the images. In the work by Meylan et al. [MDS06] the user first selects which pixels in the image can be considered highlights and then two different linear scaling functions are applied according to this classification. Rempel et al. [RTS*07] present a real-time reverse tone mapper operator (rTMO) based on a linearization of pixel values and contrast scaling, followed by a brightness enhancement function similar in spirit to the expand-map. In a series of psychophysical tests, Akyüz and co-workers [AFR*07] come up with a surprising conclusion: LDR data might not require sophisticated treatment prior to its visualization on an HDR display. By merely linearly scaling the range of the LDR input image to fit the range of the HDR display the results are considered as good as (or better than) an original HDR image. Unfortunately, they base their tests solely on correctly exposed images, and the outcome is unclear if that assumption is broken. In fact, while some of the above works present solutions to minimize noise expansion [BLDC06, RTS*07], none deal specifically with the problem of bad exposure in *imperfect*, legacy content, where the image is either under- or overexposed. Highlights in [MDS06] are in fact defined as overexposed pixels above a certain threshold value; however, the method seems to work better if these are localized to small regions of the image. It is unclear whether the algorithm would provide a pleasant solution by boosting large areas (such as an overexposed sky) the way it boosts small highlights.

We need a method to deal with imperfect content as well, but how to expand its dynamic range is not obvious. Clearly, under- and overexposure effects have been consciously used for decades, and have become standard artistic expressions, not just the result of a faulty capture process (Figure 1). Common dodge and burn techniques, for instance, are usually employed to apply local adjustments to aid tonemapping; however, they can be used for exactly the opposite reasons, to actually simulate the effects of incorrect exposure. In other words, sometimes what we call bad exposure is a deliberate decision based on artistic and aesthetic issues, and then we are facing the additional problem of carrying over the *mood* to an HDR display when reverse tone mapping is applied.

This paper aims at shedding some light onto reverse tone mapping for imperfect digital photography. We first show the results of a psychophysical test, where the subjects were presented a series of images with increasing exposures within each image set, and were asked to tag each individual image (exposure) as underexposed, correctly exposed or overexposed merely by visual inspection. We analyze the results comparing with four luminance statistics in the image: histogram, mean, median and percentage of under- and overexposed pixels. We then propose a methodology to evaluate four existing reverse tone mapping algorithms for incor-



Figure 1: Using exposure as artistic expression (Jill, by Joseph Szymanski)

rectly exposed content, also based on psychophysics. To our knowledge, this is the first time that such study is performed, and the reasons to do it are twofold: on the one hand, the fact that, as argued, a lot of the current digital content is *not* properly exposed (and complete backward compatibility is a must for HDR displays to succeed). On the other hand, before a working reverse tone mapping algorithm can be developed, it is necessary to understand all the aspects of the problem, both technical and psychophysical.

The rest of the paper is organized as follows: the next section introduces the concepts of under- and overexposure, and justifies the psychophysical approach to the following tests. In Section 3 we present the stimuli, methodology and results for our test on the perception of exposure. Section 4 explains the proposed methodology to evaluate four existing reverse tone mapping algorithms. Finally, Section 5 presents conclusions and future work.

2. Under- and overexposure

Exposure in photography can be defined as the total amount of light allowed to fall on the photographic medium during the process of taking a photograph [Kel06]. Under- or overexposure can then be loosely defined as having allowed too little or too much light. But according to what? Let us imagine the following "text-book" example: a scene made up of a green landscape, a red car and a man driving it. If the photographer wants the red car to have correct exposure then he has to measure the light reflecting off of it and sub-expose the photometer reading between one and two stops. However, if he wants the (pale) driver to be correctly exposed, he will have to over-expose one and a half stops, and if he wants the grass to be correctly-exposed he will use the exact measuring of the photometer. So, even if the camera were able to interpret such high-level components of the scene as the green landscape, the red car and the pale driver, it still could not guess the intention of the photographer.

If the images' exposure correctness could be objectively assessed using only image data (with no human interpretation), the digital cameras' firmware could in theory automatically obtain the proper exposure for every scene. Whilst most consumer cameras do offer an estimation that works well for a sufficiently large number of cases, sometimes skilled human intervention is necessary, especially at professional levels.

We thus argue that high-level semantics and human interpretation of the image are necessary in the process of determining whether an image is under- or overexposed. This is further backed by the experiments performed by Akyüz and colleagues [AFR*07]. The authors use LDR bracketed sequence as proposed in [DM97] to create the HDR images. The participants were asked to determine which single exposure was the best among the exposures used. Their results (not included in the paper, but available in [Aky]) show that participants do not always choose the image with the fewest under- or overexposed number of pixels, nor simply the middle exposure of the bracketed sequence. A high-level (and probably individual) interpretation of the scene seems to take place in the decision-making process. The design of our psychophysical tests is in part motivated by these findings.

3. Psychophysical test: exposure perception

As we have argued, under- and overexposure have apparently not yet been defined in objective terms[‡]. This suggests that there is no correlation between apparent correct exposure and objective image data, such as luminance histogram, mean, median or percentage of under- or overexposed pixels (see Figure 2), which holds for a sufficiently large number of images. It would be possible in theory to detect a subset of cases, for instance when the histogram shows null values above or below certain thresholds. But even then, false detections would happen, as in the case of low-contrast images with uniformly lit surfaces. For some applications, a useful approach may be to define a threshold under which pixels will be considered underexposed, and a second one over which overexposure is defined (which is how Meylan and colleagues define highlights in [MDS06]). However, these are operations performed at pixel level, and provide no information about the aspect of the image as a whole. More complicated cases include the possibility of an image being under- and overexposed at the same time in different areas (see Figure 3).

It thus seems that to properly classify an image as under- or overexposed we need to rely on context-dependent, high-level image semantics, as suggested in previous studies [AFR*07, Aky] and shown in Figure 2. We put this assumption to the test, by comparing the subjective perception of

[‡] This has been confirmed by interviews with professional photographers and cinematographers



Figure 2: Two different photographs with very similar luminance histogram, mean, median and percentage of saturated pixels. However, taking into account high-level semantics, the photograph on the left can be considered correctly exposed, while the one on the right is clearly overexposed.



Figure 3: Under- and overexposure in the same photograph. Not enough light reaches the corner of the wall, while there is too much light in the window area.

exposure in images with first-order image statistics: luminance histogram, mean, median and percentage of under- and overexposed pixels (defined with reference to a certain threshold). Gaining insight on this matter seems crucial for the problem of reverse tone mapping for imperfect legacy content.

3.1. Stimuli

We use images taken from 10 different scenes. The stimuli images were captured with a Nikon D200 at a resolution of 3872 by 2592 and then down-sampled to 1920 by 1080 for visualization purposes. The scenes were chosen to cover a broad range of lighting conditions and environment types. We shot a bracketed series of five exposures for each scene, ranging from clearly underexposed (labeled as 1 in the paper) to highly overexposed (labeled as 5), giving a total of 50 images used. For each scene, a tone-mapped sample is shown in Figure 4 for visualization purposes.

Luminances values for the experimental stimuli are obtained from their (R, G, B) pixel values according to $L = 0.213R + 0.715G + 0.072B$, as proposed in [RWP05]. Figure 5 shows the complete bracketed sequence for the *sunset* scene, along with the respective histograms. Tables 1 and 2 present the luminance mean and median respectively; Ta-

ble 3 shows the percentage of pixels above a given luminance threshold of 254. This value is chosen since it has been found to work well discriminating overexposed areas in photographs [RTS*07]. Finally, Table 4 shows the percentage of pixels with null luminance value, which represents our underexposed pixel threshold.

Sequence / Exposure	1	2	3	4	5
Building	126.66	158.96	183.77	211.39	232.45
Car	14.72	23.79	38.41	58.58	85.06
Indoor flower	19.67	30.29	50.28	60.41	60.51
Lake	70.77	86.47	138.04	170.17	201.54
Pencils	26.68	43.37	68.65	104.5	138.95
Computers	14.93	20.55	24.05	34.98	52.51
Waxes	29.48	50.41	78.21	111.85	148.84
Sunset	91.86	108.56	139.87	194.62	229.99
Graffiti	127.06	169.69	202.56	226.57	242.94
Strawberries	90.39	130.18	168.60	199.79	223.45

Table 1: Pixel-luminance mean for the bracketed sequence of each scene.

tinySequence / Exposure	1	2	3	4	5
Building	85	137	189	240	255
Car	6	14	27	50	86
Indoor flower	7	15	33	45	45
Lake	43	61	119	164	210
Pencils	9	18	39	83	136
Computers	0	3	5	11	26
Waxes	23	46	80	120	167
Sunset	75	97	136	218	255
Graffiti	133	189	233	255	255
Strawberries	90	140	198	246	254

Table 2: Pixel-luminance median for the bracketed sequence of each scene.

3.2. Experimental design

The design of the psychophysical experiment follows the scheme sometimes referred to as the *method of constant stimuli* [DBW08]: the fifty images are shown one by one, in random order, thus mixing both exposures and scenes. The participants are requested to classify each image in one of these groups: (1) underexposed, (2) correct, (3) overexposed. There is no fixed time for every image to be shown. The participant can move forward (to the next photograph) whenever they are done judging the current image. To ensure the

Sequence / Exposure	1	2	3	4	5
Building	4.60	19.97	40.03	44.33	47.78
Car	0.02	0.03	0.21	0.60	1.90
Indoor flower	0.64	0.79	1.30	1.95	1.98
Lake	0	0	18.71	23.93	34.49
Computers	0.14	0.54	0.90	2.32	7.36
Waxes	0	0	0	0.01	2.26
Sunset	0.01	4.45	8.35	26.16	51.72
Pencils	0	0	0	0	1.94
Graffiti	0.01	1.00	20.19	49.20	61.46
Strawberries	0	0.01	5.97	22.16	38.39

Table 3: Percentage of pixels with luminance values 254 and 255.

Sequence / Exposure	1	2	3	4	5
Building	0	0	0	0	0
Car	30.66	19.65	9.23	4.08	1.16
Indoor flower	26.11	17.10	5.99	3.62	3.50
Lake	0.10	0	0	0	0
Computers	58.06	23.53	14.11	2.70	0.12
Waxes	13.13	4.56	0.45	0	0
Sunset	0	0	0	0	0
Pencils	12.51	8.38	5.04	1.02	0.06
Graffiti	0	0	0	0	0
Strawberries	0	0	0	0	0

Table 4: Percentage of pixels with null luminance values.

validity of the data, a brief learning task is performed prior to the real test as suggested by [Ken75]: the participants are invited to judge a few images before they start classifying until they feel confident and understand the concepts. These previous images come from extra scenes are not part of the test itself. The display used was a 24-inch FP241VW model from BenQ. The experiment was set up in a darkened room in order not to reduce the perceived contrast ratio of the display (measured at 60:1). Ambient luminance measured from the wall was $26 \text{ cd}/m^2$.

A gender-balanced group of 24 participants took part in the experiment. Half of them had some photographic skills, whilst all reported normal or corrected-to-normal vision. They sat at a viewing distance of approximately a half meter from the display.

3.3. Significance of the results

Visual inspection of the results of the test (Figure 6) shows the expected logical diagonal distribution of perceived exposure. Strong backlighting of the main objects in some scenes has been mostly interpreted as under- (*indoor, car*) or overexposure (*building, sunset*), although it could be that the photographer's intention was to achieve that effect. This again indicates the need for high-level semantics and possibly human intervention when judging exposure. Some kind of machine learning or classification method, such as Support Vector Machines [Vap95] would be interesting to optimally separate images perceived as under- or overexposed, or even correctly or incorrectly (both under and over) exposed. Four of the five images with strongest gradients (the four previously mentioned plus *computers*) obtained the least number of "correct exposure" votes, suggesting that second-order statistics could provide additional insight into this topic. As expected, the histogram by itself does not provide enough information about an image's exposure.

To analyze correlations in the data, we rely on the Pearson correlation coefficient $\rho_{X,Y}$, defined as:

$$\rho_{X,Y} = \frac{E(XY) - E(X)E(Y)}{\sqrt{E(X^2) - E^2(X)} \sqrt{E(Y^2) - E^2(Y)}} \quad (1)$$

where E is the expected value operator, X is the results of

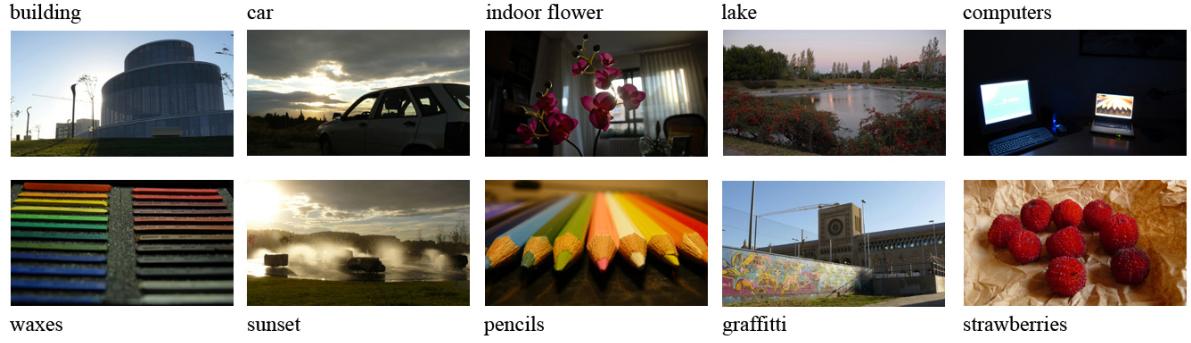


Figure 4: Tone-mapped samples of each stimuli scene.



Figure 5: The complete bracketed sequence for the sunset scene, along with the respective histograms.

the psychophysics evaluation and Y represents the objective parameter being under study (mean, media or the percentage of under- or overexposed pixels). For luminance mean and overexposure, this Pearson coefficient is $\rho_m^o = 0.869$. This is a relatively high value for psychological research, according to Cohen [Coh88]. Similar correlation exists for the luminance median and overexposure ($\rho_{md}^o = 0.846$). This correlation is logically negative for perceived underexposure but, maybe surprisingly, not so strong ($\rho_m^u = -0.726$ and $\rho_{md}^u = -0.691$).

A similar behavior can be observed for the percentage of badly exposed pixels. There is a strong positive correlation between perceived overexposure and saturated pixels ($\rho_p^o = 0.890$) but it becomes lower again for perceived underexposure and pixels with null values ($\rho_p^u = 0.675$). Although this is nothing but mere speculation at this point, these results may suggest some correlation between perceived exposure and the well-known asymmetry of the human visual system under photopic and scotopic conditions [Liv02]. We believe this is an interesting result which we plan to investigate further. Figure 7 shows these results for the case of mean and underexposure. Figure 8 shows the relation between perceived overexposure and the percentage of overexposed pixels. These two cases represent the most-correlated cases for under- and overexposure respectively. Finally, it

could be thought that perceived correct exposure may be related to the low occurrence of badly exposed pixels in the image. We found evidence of this, as indicated by its low correlation coefficient ($\rho_{sum}^c = -0.676$).

In conclusion, the two key ideas learned from this experiment, at least for the images shown and the statistics analyzed, are:

- The results seem to confirm the hypothesis that high-level semantics are needed for a proper classification of exposure. This is interesting since it apparently clashes with the notion that visual appeal is based on low-level attributes of an image [AFR*07].
- We found an asymmetry in under- and overexposure perception which may be deeply rooted in the behavior of our visual system. To confirm this, more research needs to be conducted.

4. Evaluating rTMO's with incorrect exposures

The results of the previous experiment provide us with a systematic labelling of images as under-, correctly-, and over-exposed. Given this labelling, a key question is how well the existing reverse tonemapping techniques can handle incorrectly-exposed LDR data. The aim of reverse tonemapping is to take LDR content and 'boost it' to HDR without

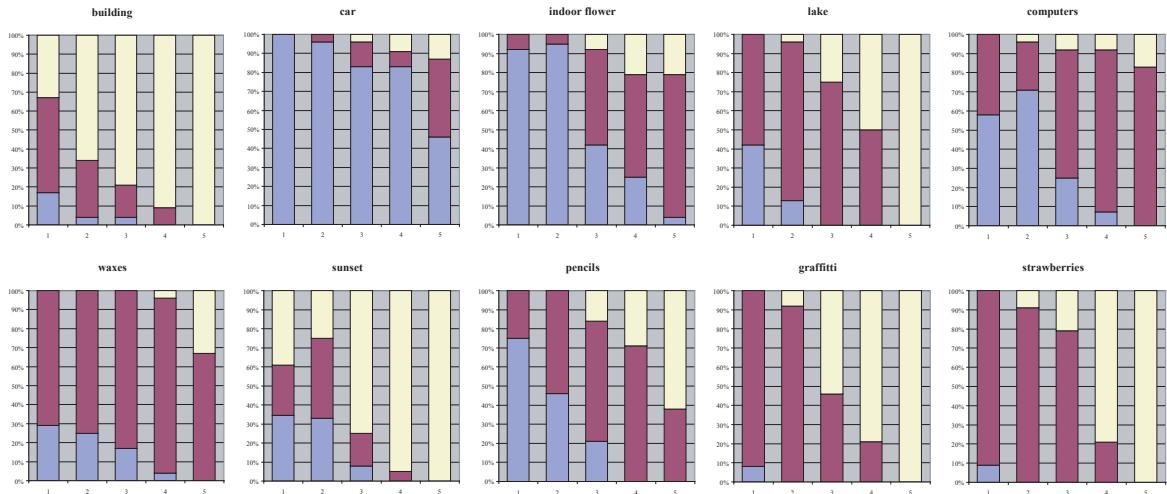


Figure 6: Results of psychophysics test: participants' stimuli taxonomy. X-axis represents the five exposures for each scene; Y-axis represents the percentage of agreement in classification (blue for underexposure, yellow for overexposure and red for correct exposure).

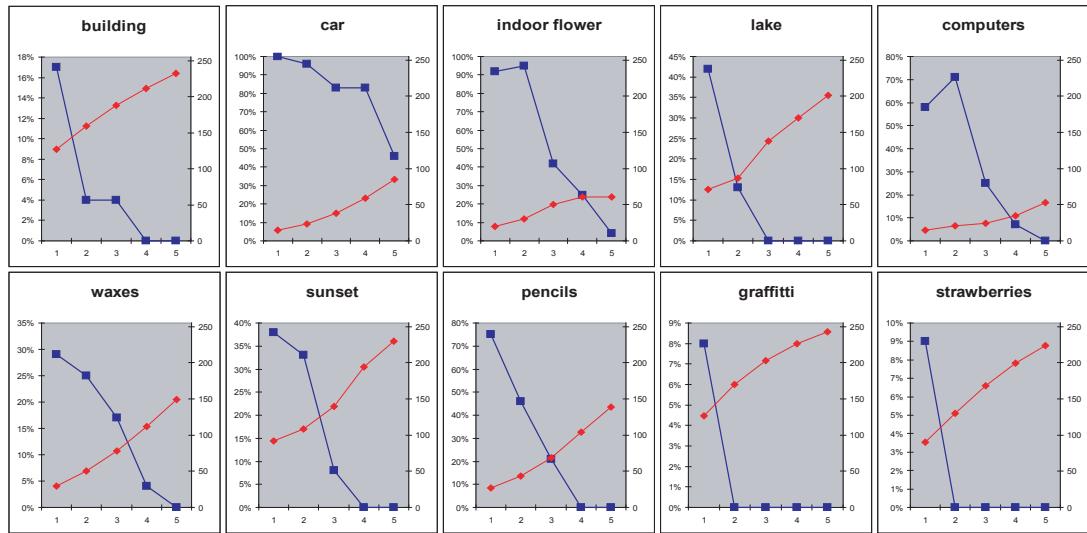


Figure 7: Inverse correlation between psychophysics results for perceived underexposure (blue) and pixel luminance mean (red). X-axis represents the five exposures for each scene; Y-axis represents the percentage of subjects who perceived the stimulus as underexposed (left) and mean luminance values (right). Note the changing scale in the Y-axis.

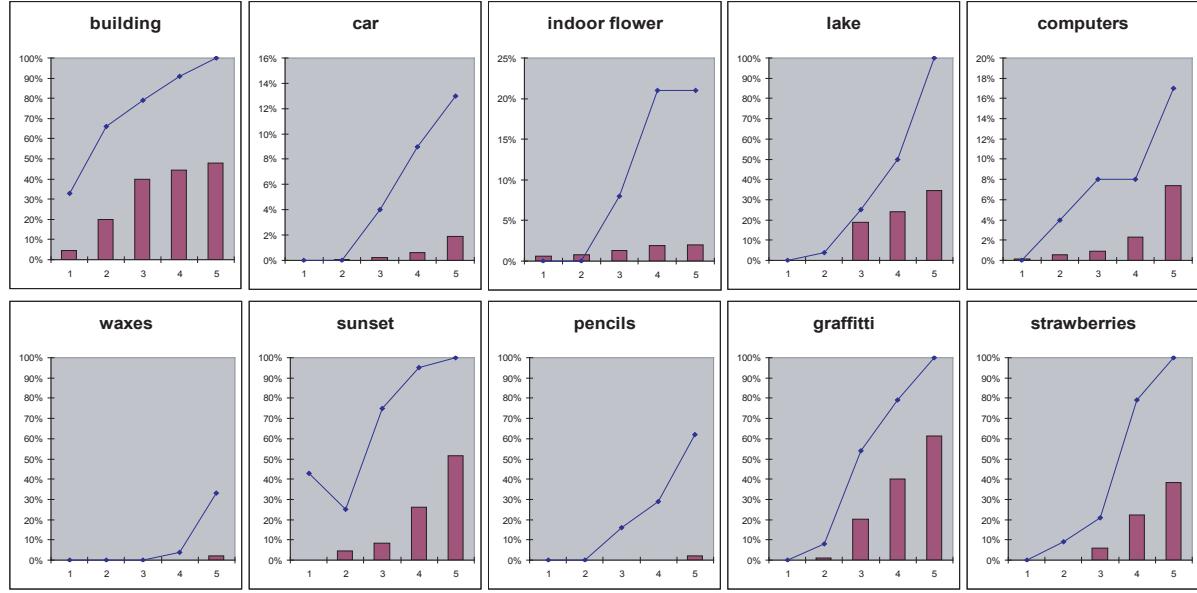


Figure 8: Correlation between psychophysics results for overexposure (blue lines) and percentage of overexposed pixels (bars). X-axis represents the five exposures for each scene; Y-axis represents the percentage of subjects who perceived the stimulus as overexposed.

introducing objectionable artifacts. Do any of the existing techniques achieve this goal? Which reverse tonemapping schemes are most appropriate for each level of exposure? To test these questions, we are currently conducting an experiment in which we ask subjects to compare the appearance of reverse tonemapped images on a Brightside DR37-P monitor. The design of the experiment is as follows.

Our goal is to perform a side-by-side comparison of the following four reverse tonemapping schemes:

1. **LDR:** the original LDR image shown on the HDR monitor,
2. **Linear:** the contrast of the original LDR image is linearly scaled to match the displayable range, as described in [AFR*07],
3. **Map:** rTMO based on expand-maps introduced by [BLDC06],
4. **Fly:** the ‘on-the-fly’ rTMO introduced by [RTS*07].

Stimuli were created as follows. For all 5 exposures of each of the 10 scenes (i.e. 50 images), we apply these four rTMOs to the image, to yield four alternative HDR renditions. On each trial, subjects are presented with the four renditions of a given image simultaneously in a randomized 2x2 grid (a ‘stimulus quadruple’). Subjects are asked to rank the four images according to how ‘visually appealing and compelling’ they appear. Subjects are instructed that this is a subjective judgment and that there is no correct answer, they should simply indicate the ordering of their personal preference. Given that previous studies showed that different judg-

ment criteria (such as ‘realism’, and ‘attractiveness’) correlate strongly [AFR*07, SH06], we decided a single subjective criterion was sufficient.

Blocks of trials consist of all 50 stimulus quadruples in pseudo-random order, with the constraint that consecutive trials cannot feature images from the same scene. Subjects are given unlimited time to respond to each trial. The entire experiment consisted of three blocks of trials. Between blocks, subjects are instructed to take a short pause before continuing with the experiment.

Once the data is analyzed, the results will provide a mean ranking score for each rTMO applied to each exposure level of each scene. This will allow us to determine which rTMO is most effective for each exposure level, and whether there is a general consensus across subjects and across scenes, or whether current rTMOs have to be selected on a case-by-case basis.

5. Conclusions and Future Work

Reverse tone mapping is a process for which no definite solution exists. With the increasing availability of HDR displays, the question of how to display the huge amount of LDR legacy content becomes an important issue. In this paper we have focused on *imperfect* legacy content, more specifically on under- and overexposed material. Rather than attempting to come up with a new reverse tone mapping algorithm, we first have looked into the crucial topic of how

exposure is perceived, so that an algorithm can be devised that keeps the look and feel of the original LDR image when viewed on an HDR display. We argue that preliminary steps in this direction are necessary, in order to avoid a proliferation in a near future of multiple co-existing rTMO's, representing partial, incomplete solutions to the problem. According to Google Scholar, there is more than 900 papers written on the topic of tone mapping, which amount to at least a few dozen different algorithms [MS08]. This is a situation we would like to avoid for reverse tone mapping.

From our psychophysical tests, two conclusions are drawn: first, the results seem to confirm that high-level semantics are probably needed for a reliable classification of exposure in images. It could be argued, though, that for some extreme cases this assumption would fail: for instance, a badly washed-out image will most likely be tagged as overexposed even in the absence of any recognizable features (and probably *due to* this absence of recognizable features). However, we believe our assumption holds for a sufficiently large number of cases. Second, we have found a clear tendency for asymmetric exposure perception, which may be related to the functioning of the human visual system.

In any case, both conclusions need to be further investigated, and in that sense we believe there is potential for lots of future research in this area. It could be argued, for instance, that the thresholds chosen for the experiments introduce bias, a topic worth looking into. We are also aware that there is an intrinsic correlation in our chosen parameters (histogram, mean, media and pixel percentages); our results should thus be seen just as a first attempt at providing a taxonomy of visual stimuli for reverse tone mapping research. Nevertheless, we hope to confirm our conclusions with additional tests which will de-correlate these parameters. Higher-order statistics will be analyzed as well, given that visual inspection of the results suggests a correlation with luminance gradients. Finally, more advanced analysis techniques need to be employed.

The psychophysical experiment proposed in Section 4 to evaluate four existing reverse tone mapping algorithms is already being performed by the authors, using a BrightSide DR37-P (display area of 32.26 by 18.15 inches, contrast ratio in excess of 200.000 : 1, black level of 0.015 cd/m² and peak luminance of 3000 cd/m²). We hope to be able to report the results soon in a subsequent publication.

6. Acknowledgements

This research has been funded by the project UZ2007-TEC06 (University of Zaragoza) and TIN2007-63025 (Spanish Ministry of Science and Technology). Olga Sorkine was partially funded by the Alexander von Humboldt Foundation, while Diego Gutierrez was additionally supported by a mobility grant by the Gobierno de Aragon (Ref: MI019/2007). The authors would like to express their gratitude to Karol Myszkowski and Matthias

Ihrke (Max Planck Institute for Informatics), for their help and advices in the early stages of this work. We also thank Francesco Banterle (University of Warwick) and Allan Rempel, Matthew Trentacoste and Wolfgang Heidrich from Dolby Canada for applying their reverse tone mapping algorithms to our stimuli, and Choss, who gave us invaluable photography information and lots of hints to guide our work.

References

- [AFR*07] AKYÜZ A. O., FLEMING R., RIECKE B. E., REINHARD E., BÜLTHOFF H. H.: Do hdr displays support ldr content?: a psychophysical evaluation. In *SIGGRAPH '07: ACM SIGGRAPH 2007 papers* (New York, NY, USA, 2007), ACM, p. 38. [1](#), [2](#), [3](#), [5](#), [7](#)
- [Aky] AKYÜZ A. O.: Homesite (http://www.coolhall.com/homepage/pubs/hdrdisp_eval/hdrdisp_project.html). [3](#)
- [BLD*07] BANTERLE F., LEDDA P., DEBATTISTA K., CHALMERS A., BLOJ M.: A framework for inverse tone mapping. *Vis. Comput.* 23, 7 (2007), 467–478. [1](#)
- [BLDC06] BANTERLE F., LEDDA P., DEBATTISTA K., CHALMERS A.: Inverse tone mapping. In *GRAPHITE '06: Proceedings of the 4th international conference on Computer graphics and interactive techniques in Australasia and Southeast Asia* (New York, NY, USA, 2006), ACM. [1](#), [2](#), [7](#)
- [Coh88] COHEN J.: *Statistical power analysis for the behavioral sciences*. Hillsdale, NJ: Lawrence Erlbaum Associates, 1988. [5](#)
- [DBW08] D. BARTZ D. CUNNINGHAM J. F., WALLRAVEN C.: The role of perception for computer graphics. *EUROGRAPHICS State of the Art Reports*, 2008. [4](#)
- [DCWP02] DEVLIN K., CHALMERS A., WILKIE A., PURGATHOFER W.: Star: Tone reproduction and physically based spectral rendering. In *State of the Art Reports, Eurographics 2002* (September 2002), Fellner D., Scopigno R., (Eds.), The Eurographics Association, pp. 101–123. [1](#)
- [DM97] DEBEVEC P. E., MALIK J.: Recovering high dynamic range radiance maps from photographs. *Computer Graphics* 31, Annual Conference Series (1997), 369–378. [3](#)
- [Kel06] KELBY S.: *The Digital Photography Book*. Peachpit Press, Berkeley, CA, USA, 2006. [2](#)
- [Ken75] KENDALL M.: *Rank Correlation Methods*. Charles Griffin & Co. Ltd, 1975. [4](#)
- [Liv02] LIVINGSTONE M.: *Vision and Art: The Biology of Seeing*. Harry N. Abrams., 2002. [5](#)
- [MDS06] MEYLAN L., DALY S., SÄIJSSTRUNK S.: The Reproduction of Specular Highlights on High Dynamic Range Displays. In *IS&T/SID 14th Color Imaging Conference* (2006). [2](#), [3](#)

- [MS08] MANTIUK R., SEIDEL H.-P.: Modeling a generic tone-mapping operator. *Computer Graphics Forum (Proceedings of Eurographics 08)* 27, 2 (2008), 699–708. 8
- [RSSF02] REINHARD E., STARK M., SHIRLEY P., FERWERDA J.: Photographic tone reproduction for digital images. *ACM Trans. Graph.* 21, 3 (2002), 267–276. 2
- [RTS*07] REMPEL A. G., TRENTACOSTE M., SEETZEN H., YOUNG H. D., HEIDRICH W., WHITEHEAD L., WARD G.: Ldr2hdr: on-the-fly reverse tone mapping of legacy video and photographs. In *SIGGRAPH '07: ACM SIGGRAPH 2007 papers* (New York, NY, USA, 2007), ACM, p. 39. 1, 2, 4, 7
- [RWP05] REINHARD E., WARD G., PATTANAIK S., DEBEVEC P.: *High Dynamic Range Imaging: Acquisition, Display and Image-Based Lighting*. Morgan Kaufmann Publishers, 2005. 1, 3
- [SH06] SEETZEN H. L. H. Y. L. W. G. W. L., HEIDRICH W.: Guidelines for contrast, brightness, and amplitude resolution of displays. In *In Society for Information Display (SID) Digest* (2006), pp. 1229–1233. 1, 7
- [SHS*04] SEETZEN H., HEIDRICH W., STUERZLINGER W., WARD G., WHITEHEAD L., TRENTACOSTE M., GHOSH A., VOROZCOVS A.: High dynamic range display systems. *Proceedings of ACM Transactions on Graphics* 23, 3 (2004), 760–768. 1
- [Vap95] VAPNIK V.: *The Nature of Statistical Learning Theory*. Springer-Verlag, Berlin, 1995. 4
- [YMMS06] YOSHIDA A., MANTIUK R., MYSZKOWSKI K., SEIDEL H.-P.: Analysis of reproducing real-world appearance on displays of varying dynamic range. In *EUROGRAPHICS 2006 (EG'06)* (Vienna, Austria, September 2006), Gröller E., Szirmay-Kalos L., (Eds.), vol. 25 of *Computer Graphics Forum*, Eurographics, Blackwell, pp. 415–426. 1